

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>4</sup> : <b>C07K 13/00, 15/00, A61K 37/00 C12P 21/00, 21/02, C12N 15/00 C07H 15/12</b>		A1	(11) International Publication Number: <b>WO 88/00205</b> (43) International Publication Date: 14 January 1988 (14.01.88)
<p>(21) International Application Number: <b>PCT/US87/01537</b></p> <p>(22) International Filing Date: 30 June 1987 (30.06.87)</p> <p>(31) Priority Application Numbers: 880,776 943,332 028,285 031,346</p> <p>(32) Priority Dates: 1 July 1986 (01.07.86) 17 December 1986 (17.12.86) 20 March 1987 (20.03.87) 26 March 1987 (26.03.87)</p> <p>(33) Priority Country: <b>US</b></p> <p>(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US);</p> <p>(72) Inventors: WANG, Elizabeth, A. ; 136 Wolf Rock Road, Carlisle, MA 01741 (US). WOZNEY, John, M. ; 59 Old Bolton Road, Hudson, MA 01749 (US). ROSEN, Vicki, A. ; 344 Marlborough Street, Apartment 4,</p>		<p>Boston, MA 02116 (US).</p> <p>(74) Agents: BAK, Mary, E. et al.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).</p> <p>(81) Designated States: AT (European patent), AU, BE (European patent), BG, BJ (OAPI patent), CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK, FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), HU, IT (European patent), JP, KR, LÜ (European patent), ML (OAPI patent), MR (OAPI patent), NL (European patent), NO, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).</p> <p>Published  <i>With international search report.</i>  <i>With amended claims.</i></p>	
<p>(54) Title: NOVEL OSTEOINDUCTIVE COMPOSITIONS</p> <p>(57) Abstract</p> <p>Human and bovine bone inductive factor products and processes. The factors may be produced by recombinant techniques and are useful in the research and treatment of bone and periodontal defects.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT Austria	FR France	ML Mali
AU Australia	GA Gabon	MR Mauritania
BB Barbados	GB United Kingdom	MW Malawi
BE Belgium	HU Hungary	NL Netherlands
BG Bulgaria	IT Italy	NO Norway
BJ Benin	JP Japan	RO Romania
BR Brazil	KP Democratic People's Republic of Korea	SD Sudan
CF Central African Republic	KR Republic of Korea	SE Sweden
CG Congo	LI Liechtenstein	SN Senegal
CH Switzerland	LK Sri Lanka	SU Soviet Union
CM Cameroon	LU Luxembourg	TD Chad
DE Germany, Federal Republic of	MC Monaco	TG Togo
DK Denmark	MG Madagascar	US United States of America
FI Finland		

## NOVEL OSTEOINDUCTIVE COMPOSITIONS

The present invention relates to novel proteins and processes for obtaining them. These proteins are capable of inducing cartilage and bone formation.

Background

Bone is a highly specialized tissue characterized by an extensive matrix structure formed of fibrous bundles of the protein collagen, and proteoglycans, noncollagenous proteins, lipids and acidic proteins. The processes of bone formation and renewal/repair of bone tissue, which occur continuously throughout life, are performed by specialized cells. Normal embryonic long bone development is preceded by formation of a cartilage model. Bone growth is presumably mediated by "osteoblasts" (bone-forming cells), while remodeling of bone is apparently accomplished by the joint activities of bone-resorbing cells, called "osteoclasts" and osteoblasts. A variety of osteogenic, cartilage-inducing and bone inducing factors have been described. See, e.g. European patent applications 148,155 and 169,016 for discussions thereof.

Brief Description of the Invention

The present invention provides novel proteins in purified form. Specifically, four of the novel proteins are designated BMP-1, BMP-2 Class I (or BMP-2), BMP-3, and BMP-2 Class II (or BMP-4) wherein BMP is bone morphogenic protein. These proteins are characterized by peptide sequences the same as or substantially homologous to amino acid sequences illustrated in Tables II through VIII below. They are capable of inducing bone formation at a predetermined site. These bone inductive factors are further characterized by biochemical and biological characteristics including activity at a concentration of 10 to 1000ng/gram of bone in an *in vivo* rat bone formation assay described below. Proteins of this invention may be encoded by the DNA sequences depicted in the Tables or by sequences capable

of hybridizing thereto and coding for polypeptides with bone growth factor biological properties or other variously modified sequences demonstrating such properties.

One of the proteins of the invention is designated BMP-1. A portion of the human BMP-1 or hBMP-1 is characterized by the same or substantially the same peptide sequence as that of amino acid #1 through amino acid #37 of Table V, below which represents a genomic hBMP-1 fragment or amino acid #1 through amino acid #730 of Table VI which represents the hBMP-1 cDNA. hBMP-1 or a related bone inductive factor may be further characterized by at least a portion of these sequences. These peptide sequences are encoded by the same or substantially the same DNA sequence, as depicted in nucleotide #3440 through nucleotide #3550 of Table V and in nucleotide #36 through nucleotide #2225 of Table VI, respectively. These hBMP-1 polypeptides are further characterized by the ability to induce bone formation. hBMP-1 demonstrates activity in an in vivo rat bone formation assay at a concentration of 10 to 1000ng/gram of bone.

The homologous bovine growth factor of the invention, designated bBMP-1, is characterized by a peptide sequence containing the same or substantially the same sequence as that of amino acid #1 through amino acid #37 of Table II below which represents a genomic bBMP-1 fragment. This peptide sequence is encoded by the same or substantially the same DNA sequence as depicted in nucleotide #294 through nucleotide #404 of Table II. The bovine peptide sequence identified in Table II below is also 37 amino acids in length. bBMP-1 is further characterized by the ability to induce bone formation.

Another bone inductive protein composition of the invention is designated BMP-2 Class I (or BMP-2). It is characterized by at least a portion of a peptide sequence the same or substantially the same as that of amino acid #1 through amino acid #396 of Table VII which represents the cDNA hBMP-2 Class I. This peptide sequence is encoded by the same or

substantially the same DNA sequence, as depicted in nucleotide #356 through nucleotide #1543 of Table VII. The human peptide sequence identified in Table VII is 396 amino acids in length. hBMP-2 or related bone inductive proteins may also be characterized by at least a portion of this peptide sequence. hBMP-2 Class I is further characterized by the ability to induce bone formation.

The homologous bovine bone inductive protein of the invention designated bBMP-2 Class I (or bBMP-2), has a DNA sequence identified in Table III below which represents the genomic sequence. This bovine DNA sequence has a prospective 129 amino acid coding sequence followed by approximately 205 nucleotides (a presumptive 3' non-coding sequence). bBMP-2, Class I is further characterized by the ability to induce bone formation. A further bone inductive protein composition of the invention is designated BMP-2 Class II or BMP-4. The human protein hBMP-2 Class II (or hBMP-4) is characterized by at least a portion of the same or substantially the same peptide sequence between amino acid #1 through amino acid #408 of Table VIII, which represents the cDNA of hBMP-2 Class II. This peptide sequence is encoded by at least a portion of the same or substantially the same DNA sequence as depicted in nucleotide #403 through nucleotide #1626 of Table VIII. This factor is further characterized by the ability to induce bone formation.

Still another bone inductive factor of the invention, BMP-3, is represented by the bovine homolog bBMP-3. bBMP-3 is characterized by the DNA sequence and amino acid sequence of Table IV A and B which represents the bovine genomic sequence. It is characterized by at least a portion of a peptide sequence the same or substantially the same as amino acid #1 through amino acid #175 of Table IV A and B. BMP-3 is further characterized by the ability to induce bone formation. The bovine factor may be employed as a tool for obtaining the analogous human BMP-3 protein or other mammalian bone inductive proteins. The proper characterization of this bovine bone

inductive factor provides the essential "starting point" for the method employing this sequence. The method, employing techniques known to those skilled in the art of genetic engineering, involves using the bovine DNA sequence as a probe to screen a human genomic or cDNA library; and identifying the DNA sequences which hybridize to the probes. A clone with a hybridizable sequence is plaque purified and the DNA isolated therefrom, subcloned and subjected to DNA sequence analysis. Thus as another aspect of this invention is a human protein hBMP-3, produced by this method.

Another aspect of the invention provides pharmaceutical compositions containing a therapeutically effective amount of one or more bone growth factor polypeptides according to the invention in a pharmaceutically acceptable vehicle. These compositions may further include other therapeutically useful agents. They may also include an appropriate matrix for delivering the proteins to the site of the bone defect and for providing a structure for bone growth. These compositions may be employed in methods for treating a number of bone defects and periodontal disease. These methods, according to the invention, entail administering to a patient needing such bone formation an effective amount of at least one of the novel proteins BMP-1, BMP-2 Class I, BMP-2 Class-II, and BMP-3 as described herein.

Still a further aspect of the invention are DNA sequences coding on expression for a human or bovine polypeptide having the ability to induce bone formation. Such sequences include the sequence of nucleotides in a 5' to 3' direction illustrated in Tables II through VIII. Alternatively, a DNA sequence which hybridizes under stringent conditions with the DNA sequences of Tables II - VIII or a DNA sequence which hybridizes under non-stringent conditions with the illustrated DNA sequences and which codes on expression for a protein having at least one bone growth factor biological property are included in the present invention. Finally, allelic or other variations of the

sequences of Tables II through VIII, whether such nucleotide changes result in changes in the peptide sequence or not, are also included in the present invention.

Still a further aspect of the invention is a vector containing a DNA sequence as described above in operative association with an expression control sequence. Such vector may be employed in a novel process for producing a bone growth factor polypeptide in which a cell line transformed with a DNA sequence encoding expression of a bone growth factor polypeptide in operative association with an expression control sequence therefor, is cultured. This claimed process may employ a number of known cells as host cells for expression of the polypeptide. Presently preferred cell lines are mammalian cell lines and bacterial cells.

Other aspects and advantages of the present invention will be apparent upon consideration of the following detailed description and preferred embodiments thereof.

#### Detailed Description of the Invention

The proteins of the present invention are characterized by amino acid sequences or portions thereof the same as or substantially homologous to the sequences shown in Tables II - VIII below. These proteins are also characterized by the ability to induce bone formation.

The bone growth factors provided herein also include factors encoded by the sequences similar to those of Tables II - VIII, but into which modifications are naturally provided (e.g. allelic variations in the nucleotide sequence which may result in amino acid changes in the polypeptide) or deliberately engineered. For example, synthetic polypeptides may wholly or partially duplicate continuous sequences of the amino acid residues of Tables II - VIII. These sequences, by virtue of sharing primary, secondary, or tertiary structural and conformational characteristics with bone growth factor polypeptides of Tables II - VIII may possess bone growth factor biological properties in common therewith. Thus, they may be

employed as biologically active substitutes for naturally-occurring bone growth factor polypeptides in therapeutic processes.

Other specific mutations of the sequences of the bone growth factors described herein involve modifications of one or both of the glycosylation sites. The absence of glycosylation or only partial glycosylation results from amino acid substitution or deletion at one or both of the asparagine-linked glycosylation recognition sites present in the sequences of the bone growth factors shown in Tables II-VIII. The asparagine-linked glycosylation recognition sites comprise tripeptide sequences which are specifically recognized by appropriate cellular glycosylation enzymes. These tripeptide sequences are either asparagine-X-threonine or asparagine-X-serine, where X is usually any amino acid. A variety of amino acid substitutions or deletions at one or both of the first or third amino acid positions of a glycosylation recognition site (and/or amino acid deletion at the second position) results in non-glycosylation at the modified tripeptide sequence.

The present invention also encompasses the novel DNA sequences, free of association with DNA sequences encoding other proteinaceous materials, and coding on expression for bone growth factors. These DNA sequences include those depicted in Tables II - VIII in a 5' to 3' direction and those sequences which hybridize under stringent hybridization conditions [see, T. Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), pages 387 to 389] to the DNA sequences of Tables II - VIII.

DNA sequences which hybridize to the sequences of Tables II - VIII under relaxed hybridization conditions and which code on expression for bone growth factors having bone growth factor biological properties also encode bone growth factors of the invention. For example, a DNA sequence which shares regions of significant homology, e.g., sites of glycosylation

or disulfide linkages, with the sequences of Tables II - VIII and encodes a bone growth factor having one or more bone growth factor biological properties clearly encodes a member of this novel family of growth factors, even if such a DNA sequence would not stringently hybridize to the sequence of Tables II - VIII.

Similarly, DNA sequences which code for bone growth factor polypeptides coded for by the sequences of Tables II - VIII, but which differ in codon sequence due to the degeneracies of the genetic code or allelic variations (naturally-occurring base changes in the species population which may or may not result in an amino acid change) also encode the novel growth factors described herein. Variations in the DNA sequences of Tables II - VIII which are caused by point mutations or by induced modifications to enhance the activity, half-life or production of the polypeptides encoded thereby are also encompassed in the invention.

Another aspect of the present invention provides a novel method for producing the novel osteoinductive factors. The method of the present invention involves culturing a suitable cell or cell line, which has been transformed with a DNA sequence coding on expression for a novel bone growth factor polypeptide of the invention, under the control of known regulatory sequences. Suitable cells or cell lines may be mammalian cells, such as Chinese hamster ovary cells (CHO). The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. See, e.g., Gething and Sambrook, Nature, 293:620-625 (1981), or alternatively, Kaufman et al, Mol. Cell. Biol., 5(7):1750-1759 (1985) or Howley et al, U.S. Patent 4,419,446. Another suitable mammalian cell line, which is described in the accompanying examples, is the monkey COS-1 cell line. A similarly useful mammalian cell line is the CV-1 cell line.

Bacterial cells are suitable hosts. For example, the

various strains of E. coli (e.g., HB101, MC1061) are well-known as host cells in the field of biotechnology. Various strains of B. subtilis, Pseudomonas, other bacilli and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for expression of the polypeptides of the present invention. Additionally, where desired, insect cells may be utilized as host cells in the method of the present invention. See, e.g. Miller et al, Genetic Engineering, 8:277-298 (Plenum Press 1986) and references cited therein.

Another aspect of the present invention provides vectors for use in the method of expression of these novel osteoinductive polypeptides. Preferably the vectors contain the full novel DNA sequences described above which code for the novel factors of the invention. Additionally the vectors also contain appropriate expression control sequences permitting expression of the bone inductive protein sequences. Alternatively, vectors incorporating modified sequences as described above are also embodiments of the present invention and useful in the production of the bone inductive proteins. The vectors may be employed in the method of transforming cell lines and contain selected regulatory sequences in operative association with the DNA coding sequences of the invention which are capable of directing the replication and expression thereof in selected host cells. Useful regulatory sequences for such vectors are known to one of skill in the art and may be selected depending upon the selected host cells. Such selection is routine and does not form part of the present invention.

A protein of the present invention, which induces bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures. An osteogenic preparation employing one or more of the proteins of the invention may have prophylactic use in closed as well as open

fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery. An osteogenic factor of the invention may be valuable in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. Of course, the proteins of the invention may have other therapeutic uses.

A further aspect of the invention is a therapeutic method and composition for repairing fractures and other conditions related to bone defects or periodontal diseases. Such a composition comprises a therapeutically effective amount of at least one of the bone inductive factor proteins of the invention. The bone inductive factors according to the present invention may be present in a therapeutic composition in admixture with a pharmaceutically acceptable vehicle or matrix. Further therapeutic methods and compositions of the invention comprise a therapeutic amount of a bone inductive factor of the invention with a therapeutic amount of at least one of the other bone inductive factors of the invention. Additionally, the proteins according to the present invention or a combination of the proteins of the present invention may be co-administered with one or more different osteoinductive factors with which it may interact. Further, the bone inductive proteins may be combined with other agents beneficial to the treatment of the bone defect in question. Such agents include, but are not limited to various growth factors. The preparation of such physiologically acceptable protein compositions, having due regard to pH, isotonicity, stability and the like, is within the skill of the art.

In particular, BMP-1 may be used individually in a

composition. BMP-1 may also be used in combination with one or more of the other proteins of the invention. BMP-1 and BMP-2 Class I may be used in combination. BMP-1 and BMP-2 Class II may also be used in combination. BMP-1 and BMP-3 may be used in combination. Furthermore, BMP-1 may be used in combination with two or three of the other proteins of the invention. For example, BMP-1, BMP-2 Class I, and BMP-2 Class II may be combined. BMP-1 may also be combined with BMP-2 Class I, and BMP-3. Further, BMP-1 may be combined with BMP-2 Class II, and BMP-3. BMP-1, BMP-2 Class I, BMP-2 Class II, and BMP-3 may be combined.

BMP-2 Class I may be used individually in a pharmaceutical composition. BMP-2 Class I may also be used in combination with one or more of the other proteins of the invention. BMP-2 Class I may be combined with BMP-2 Class II. It may also be combined with BMP-3. Further BMP-2 Class I may be combined with BMP-2 Class II and BMP-3.

BMP-2 Class II may be used individually in pharmaceutical composition. In addition, it may be used in combination with other proteins as identified above. Further it may be used in combination with BMP-3.

BMP-3 may be used individually in a composition. It may further be used in the various combinations identified above.

The therapeutic method includes locally administering the composition as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone damage. Preferably, the bone growth inductive factor composition would include a matrix capable of delivering the bone inductive factor to the site of bone damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of other materials presently in use for other implanted medical

applications.

The choice of material is based on, for example, biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. Similarly, the application of the osteoinductive factors will define the appropriate formulation. Potential matrices for the osteoinductive factors may be biodegradable and chemically defined, such as, but not limited to calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyanhydrides; biodegradable and biologically well defined, such as bone or dermal collagen, other pure proteins or extracellular matrix components; nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics; or combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics might also be altered in composition, such as in calcium-aluminate-phosphate and processing to alter for example, pore size, particle size, particle shape, and biodegradability.

The dosage regimen will be determined by the attending physician considering various factors which modify the action of such a growth factor, e.g. amount of bone weight desired to be formed, the site of bone damage, the condition of the damaged bone, the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and the composition of BMP's. The addition of other known growth factors, such as IGF 1 (insulin like growth factor 1), to the final composition, may also effect the dosage. Generally, the dosage regimen should be in the range of approximately 10 to  $10^6$  nanograms of protein per gram of bone weight desired. Progress can be monitored by periodic assessment of bone growth and/or repair, e.g. x-rays. Such therapeutic compositions are also presently valuable for veterinary applications due to the lack of species specificity

in bone inductive factors. Particularly domestic animals and thoroughbred horses in addition to humans are desired patients for such treatment with the bone inductive factors of the present invention.

The following examples illustrate practice of the present invention in recovering and characterizing the bovine proteins and employing them to recover the human proteins, obtaining the human proteins and in expressing the proteins via recombinant techniques.

#### EXAMPLE I

##### Isolation of Bovine Bone Inductive Factor

Ground bovine bone powder (20-120 mesh, Helitrex) is prepared according to the procedures of M. R. Urist et al., Proc. Natl Acad. Sci USA, 70:3511 (1973) with elimination of some extraction steps as identified below. Ten kgs of the ground powder is demineralized in successive changes of 0.6N HCl at 4°C over a 48 hour period with vigorous stirring. The resulting suspension is extracted for 16 hours at 4°C with 50 liters of 2M CaCl<sub>2</sub> and 10mM ethylenediamine-tetraacetic acid [EDTA], and followed by extraction for 4 hours in 50 liters of 0.5M EDTA. The residue is washed three times with distilled water before its resuspension in 20 liters of 4M guanidine hydrochloride [GuCl], 20mM Tris (pH 7.4), 1mM N-ethylmaleimide, 1mM iodoacetamide, 1mM phenylmethylsulfonyl fluorine as described in Clin. Orthop. Rel. Res., 171: 213 (1982). After 16 to 20 hours the supernatant is removed and replaced with another 10 liters of GuCl buffer. The residue is extracted for another 24 hours.

The crude GuCl extracts are combined, concentrated approximately 20 times on a Pellicon apparatus with a 10,000 molecular weight cut-off membrane, and then dialyzed in 50mM Tris, 0.1M NaCl, 6M urea (pH7.2), the starting buffer for the first column. After extensive dialysis the protein is loaded on a 4 liter DEAE cellulose column and the unbound fractions

are collected.

The unbound fractions are concentrated and dialyzed against 50mM NaAc, 50mM NaCl (pH 4.6) in 6M urea. The unbound fractions are applied to a carboxymethyl cellulose column. Protein not bound to the column is removed by extensive washing with starting buffer, and the bone inductive factor containing material desorbed from the column by 50mM NaAc, 0.25mM NaCl, 6M urea (pH 4.6). The protein from this step elution is concentrated 20- to 40- fold, then diluted 5 times with 80mM KPO<sub>4</sub>, 6M urea (pH6.0). The pH of the solution is adjusted to 6.0 with 500mM K<sub>2</sub>HPO<sub>4</sub>. The sample is applied to an hydroxylapatite column (LKB) equilibrated in 80mM KPO<sub>4</sub>, 6M urea (pH6.0) and all unbound protein is removed by washing the column with the same buffer. Bone inductive factor activity is eluted with 100mM KPO<sub>4</sub> (pH7.4) and 6M urea.

The protein is concentrated approximately 10 times, and solid NaCl added to a final concentration of 0.15M. This material is applied to a heparin - Sepharose column equilibrated in 50mM KPO<sub>4</sub>, 150mM NaCl, 6M urea (pH7.4). After extensive washing of the column with starting buffer, a protein with bone inductive factor activity is eluted by 50mM KPO<sub>4</sub>, 700mM NaCl, 6M urea (pH7.4). This fraction is concentrated to a minimum volume, and 0.4ml aliquots are applied to Superose 6 and Superose 12 columns connected in series, equilibrated with 4M GuCl, 20mM Tris (pH7.2) and the columns developed at a flow rate of 0.25ml/min. The protein demonstrating bone inductive factor activity has a relative migration corresponding to approximately 30,000 dalton protein.

The above fractions are pooled, dialyzed against 50mM NaAc, 6M urea (pH4.6), and applied to a Pharmacia MonoS HR column. The column is developed with a gradient to 1.0M NaCl, 50mM NaAc, 6M urea (pH4.6). Active fractions are pooled and brought to pH3.0 with 10% trifluoroacetic acid (TFA). The material is applied to a 0.46 x 25cm Vydac C4 column in 0.1% TFA and the column developed with a gradient to 90% acetonitrile, 0.1% TFA

(31.5% acetonitrile, 0.1% TFA to 49.5% acetonitrile, 0.1% TFA in 60 minutes at 1ml per minute). Active material is eluted at approximately 40-44% acetonitrile. Aliquots of the appropriate fractions are iodinated by one of the following methods: P. J. McConahey et al, Int. Arch. Allergy, 29:185-189 (1966); A. E. Bolton et al, Biochem J., 133:529 (1973); and D. F. Bowen-Pope, J. Biol. Chem., 237:5161 (1982). The iodinated proteins present in these fractions are analyzed by SDS gel electrophoresis and urea Triton X 100 isoelectric focusing. At this stage, the bone inductive factor is estimated to be approximately 10-50% pure.

#### EXAMPLE II

##### Characterization of Bovine Bone Inductive Factor

###### A. Molecular Weight

Approximately 20ug protein from Example I is lyophilized and redissolved in 1X SDS sample buffer. After 15 minutes of heating at 37°C, the sample is applied to a 15% SDS polyacrylamide gel and then electrophoresed with cooling. The molecular weight is determined relative to prestained molecular weight standards (Bethesda Research Labs). Immediately after completion, the gel lane containing bone inductive factor is sliced into 0.3cm pieces. Each piece is mashed and 1.4ml of 0.1% SDS is added. The samples are shaken gently overnight at room temperature to elute the protein. Each gel slice is desalted to prevent interference in the biological assay. The supernatant from each sample is acidified to pH 3.0 with 10% TFA, filtered through a 0.45 micron membrane and loaded on a 0.46cm x 5cm C4 Vydac column developed with a gradient of 0.1% TFA to 0.1% TFA, 90% CH<sub>3</sub>CN. The appropriate bone inductive factor - containing fractions are pooled and reconstituted with 20mg rat matrix. In this gel system, the majority of bone inductive factor fractions have the mobility of a protein having a molecular weight of approximately 28,000 - 30,000 daltons.

### B. Isoelectric Focusing

The isoelectric point of bone inductive factor activity is determined in a denaturing isoelectric focusing system. The Triton X100 urea gel system (Hoeffer Scientific) is modified as follows: 1) 40% of the ampholytes used are Servalyte 3/10; 60% are Servalyte 7-9. 2) The catholyte used is 40mM NaOH. Approximately 20ug of protein from Example I is lyophilized, dissolved in sample buffer and applied to the isoelectrofocusing gel. The gel is run at 20 watts, 10°C for approximately 3 hours. At completion the lane containing bone inductive factor is sliced into 0.5 cm slices. Each piece is mashed in 1.0ml 6M urea, 5mM Tris (pH 7.8) and the samples agitated at room temperature. The samples are acidified, filtered, desalted and assayed as described above. The major portion of activity as determined in the assay described in Example III migrates in a manner consistent with a pI of 8.8 - 9.2.

### C. Subunit Characterization

The subunit composition of bone inductive factor is also determined. Pure bone inductive factor is isolated from a preparative 15% SDS gel as described above. A portion of the sample is then reduced with 5mM DTT in sample buffer and re-electrophoresed on a 15% SDS gel. The approximately 30kd protein yields two major bands at approximately 20kd and 18kd, as well as a minor band at 30kd. The broadness of the two bands indicates heterogeneity caused most probably by glycosylation, other post translational modification, proteolytic degradation or carbamylation.

### EXAMPLE III

#### Biological Activity of Bone Inductive Factor

A rat bone formation assay according to the general procedure of Sampath and Reddi, Proc. Natl. Acad. Sci. U.S.A.,

80:6591-6595. (1983) is used to evaluate the osteogenic activity of the bovine bone inductive factor of the present invention obtained in Example I. This assay can also be used to evaluate bone inductive factors of other species. The ethanol precipitation step is replaced by dialyzing the fraction to be assayed against water. The solution or suspension is then redissolved in a volatile solvent, e.g. 0.1 - 0.2 % TFA, and the resulting solution added to 20mg of rat matrix. This material is frozen and lyophilized and the resulting powder enclosed in #5 gelatin capsules. The capsules are implanted subcutaneously in the abdominal thoracic area of 21 - 49 day old male long Evans rats. The implants are removed after 7 - 14 days. Half of each implant is used for alkaline phosphatase analysis [See, A. H. Reddi et al., Proc. Natl Acad Sci., 69:1601 (1972)] and half is fixed and processed for histological analysis. Routinely, lum glycolmethacrylate sections are stained with Von Kossa and acid fuchsin to detect new bone mineral. Alkaline phosphatase, an enzyme produced by chondroblasts and osteoblasts in the process of matrix formation, is also measured. New cartilage and bone formation often correlates with alkaline phosphatase levels. Table I below illustrates the dose response of the rat matrix samples including a control not treated with bone inductive factor.

TABLE I

<u>Protein*</u>	<u>Implanted ug</u>	<u>Cartilage</u>	<u>Alk. Phos.u/l</u>
	7.5	2	Not done
	2.5	3	445.7
	0.83	3	77.4
	0.28	0	32.5
	0.00	0	31.0

\*At this stage the bone inductive factor is approximately 10-15% pure.

The bone or cartilage formed is physically confined to the space occupied by the matrix. Samples are also analyzed by SDS gel electrophoresis and isoelectric focusing as described

above, followed by autoradiography. Analysis reveals a correlation of activity with protein bands at 28 - 30kd and a pI 9.0. An extinction coefficient of 1 OD/mg-cm is used as an estimate for protein and approximating the purity of bone inductive factor in a particular fraction. In the in vivo rat bone formation assays on dilutions as described above, the protein is active in vivo at 10 to 200ng protein/gram bone to probably greater than 1ug protein/gram bone.

#### EXAMPLE IV

##### Bovine Bone Inductive Factor Protein Composition

The protein composition of Example IIIA of molecular weight 28 - 30kd is reduced as described in Example IIC and digested with trypsin. Eight tryptic fragments are isolated by standard procedures having the following amino acid sequences:

Fragment 1: A A F L G D I A L D E E D L G

Fragment 2: A F Q V Q Q A A D L

Fragment 3: N Y Q D M V V E G

Fragment 4: S T P A Q D V S R

Fragment 5: N Q E A L R

Fragment 6: L S E P D P S H T L E E

Fragment 7: F D A Y Y

Fragment 8: L K P S N ? A T I Q S I V E

A less highly purified preparation of protein from bovine bone is prepared according to a purification scheme similar to that described in Example I. The purification basically varies from that previously described by omission of the DE-52 column, the CM cellulose column and the mono S column, as well as a reversal in the order of the hydroxylapatite and heparin sepharose columns. Briefly, the concentrated crude 4 M extract is brought to 85% final concentration of ethanol at 4 degrees. The mixture is then centrifuged, and the precipitate redissolved in 50 mM Tris, 0.15 M NaCl, 6.0 M urea. This material is then fractionated on Heparin Sepharose as described. The Heparin bound material

is fractionated on hydroxyapatite as described. The active fractions are pooled, concentrated, and fractionated on a high resolution gel filtration (TSK 30000 in 6 M guanidinium chloride, 50 mM Tris, pH 7.2). The active fractions are pooled, dialyzed against 0.1% TFA, and then fractionated on a C4 Vydac reverse phase column as described. The preparation is reduced and electrophoresed on an acrylamide gel. The protein corresponding to the 18K band is eluted and digested with trypsin. Tryptic fragments are isolated having the following amino acid sequences:

Fragment 9: S L K P S N H A T I Q S ? V

Fragment 10: S F D A Y Y C S ? A

Fragment 11: V Y P N M T V E S C A

Fragment 12: V D F A D I ? W

Tryptic Fragments 7 and 8 are noted to be substantially the same as Fragments 10 and 9, respectively.

A. bBMP-1

Probes consisting of pools of oligonucleotides (or unique oligonucleotides) are designed according to the method of R. Lathe, *J. Mol. Biol.*, 183 (1):1-12 (1985) and synthesized on an automated DNA synthesizer. One probe consists of a relatively long (32 nucleotides) "guessmer" [See J. J. Toole et al, *Nature*, 312:342-347 (1984)] of the following nucleotide sequence:

TCCTCATCCAGGGCAATGTCGCCAGGAAGGC

Because the genetic code is degenerate (more than one codon can code for the same amino acid), the number of oligonucleotides in a probe pool is reduced based on the frequency of codon usage in eukaryotes, the relative stability of G:T base pairs, and the relative infrequency of the dinucleotide CpG in eukaryotic coding sequences [see Toole et al., supra.]. The second set of probes consists of shorter oligonucleotides (17 nucleotides in length) which contain all possible sequences that could encode the amino acids. The second set of probes has the following sequences:

(a) A [A/G] [A/G] TC [T/C] TC [T/C] TC [A/G] TC [T/C] AA

(b) A [A/G] [A/G] TC [T/C] TC [T/C] TC [A/G] TCNAG

Bracketed nucleotides are alternatives. "N" means either A, T, C or G.

In both cases the regions of the amino acid sequence used for probe design are chosen by avoiding highly degenerate codons where possible. The oligonucleotides are synthesized on an automated DNA synthesizer; the probes are then radioactively labeled with polynucleotide kinase and  $^{32}\text{P}$ -ATP.

These two sets of probes are used to screen a bovine genomic recombinant library. The library is constructed as follows: Bovine liver DNA is partially digested with the restriction endonuclease enzyme Sau 3A and sedimented through a sucrose gradient. Size fractionated DNA in the range of 15-30kb is then ligated to the bacteriophage Bam HI vector EMBL3 [Frischauf et al, J. Mol. Biol., 170:827-842 (1983)]. The library is plated at 8000 recombinants per plate. Duplicate nitrocellulose replicas of the plaques are made and amplified according to a modification of the procedure of Woo et al, Proc. Natl. Acad. Sci. USA, 75:3688-91 (1978).

The 32 mer probe is kinased with  $^{32}\text{P}$ -gamma-ATP and hybridized to one set of filters in 5X SSC, 0.1% SDS, 5X Denhardts, 100ug/ml salmon sperm DNA at 45 degrees C and washed with 5X SSC, 0.1% SDS at 45 degrees C. The 17 mer probes are kinased and hybridized to the other set of filters in 3M tetramethylammonium chloride (TMAC), 0.1M sodium phosphate pH6.5, 1mM EDTA, 5X Denhardts, 0.6% SDS, 100ug/ml salmon sperm DNA at 48 degrees C, and washed in 3M TMAC, 50mM Tris pH8.0 at 50 degrees C. These conditions minimize the detection of mismatches to the 17 mer probe pool [see, Wood et al, Proc. Natl. Acad. Sci. U.S.A., 82:1585-1588 (1985)]. 400,000 recombinants are screened by this procedure and one duplicate positive is plaque purified. DNA is isolated from a plate lysate of this recombinant bacteriophage designated lambda bP-50. bP-50 was deposit d December 16, 1986 with the American

Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland USA (hereinafter the "ATCC") under accession number 40295.. This deposit as well as the other deposits contained herein meets the requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulations thereunder. This bp-50 clone encodes at least a portion of the bovine bone growth factor designated bBMP-1.

The oligonucleotide hybridizing region of this bBMP-1 clone is localized to an approximately 800bp Eco RI fragment which is subcloned into M13 and sequenced by standard techniques. The partial DNA sequence and derived amino acid sequence of lambda bp-50 are shown below in Table II. The amino acid sequences corresponding to the tryptic fragments isolated from the bovine bone 28 to 30kd material are underlined in Table II. The first underlined portion of the sequence corresponds to tryptic Fragment 1 above from which the oligonucleotide probes are designed. The second underlined portion corresponds to tryptic Fragment 2 above. The predicted amino acid sequence indicates that tryptic Fragment 2 is preceded by a basic residue (R) as expected considering the specificity of trypsin. The nucleic acid sequence preceding the couplet CT at nucleotide positions #292-293 in Table II is presumed to be an intron (noncoding sequence) based on the presence of a consensus acceptor sequence (i.e., a pyrimidine rich tract, TCTCTCTCC, followed by AG) and the lack of a basic residue in the appropriate position of the derived amino acid sequence. This bBMP-1 genomic sequence appears in Table II. The presumptive bBMP-1 peptide sequence from this genomic clone is 37 amino acids in length and is encoded by the DNA sequence from nucleotide #294 through #404 in Table II.

TABLE II

280            290            (1)            308            323  
CCATGCGCTCT TCTCTCTCCA GCT GCC TTC CTT GGG GAC ATC GCC CTG GAC GAG GAG  
Ala Phe Leu Gly Asp Ile Ala Leu Asp Glu Glu

338            353            368  
GAC TTG AGG GCC TTC CAA GTG CAG CAG GCT GCG GAC CTC AGA CAG CGT GCA ACC  
Asp Leu Arg Ala Phe Gln Val Gln Gln Ala Ala Asp Leu Arg Gln Arg Ala Thr

383            398            (37)            414            424  
CGC AGG TCT TCC ATC AAA GCT GCA GGTACACTGG GTACAGGCGCA  
Arg Arg Ser Ser Ile Lys Ala Ala

**B. bBMP-2**

Two probes consisting of pools of oligonucleotides are designed on the basis of the amino acid sequence of Fragment 3 and synthesized on an automated DNA synthesizer as described above.

Probe #1: A C N A C C A T [A/G] T C [T/C] T G [A/G] A T

Probe #2: C A [A/G] G A [T/C] A T G G T N G T N G A

These probes are radioactively labeled and employed to screen the bovine genomic library constructed as described in part A except that the vector is lambda J1 Bam H1 arms [Mullins et al Nature 308: 856-858 (1984).] The radioactively labelled 17-mer Probe #1 is hybridized to the set of filters according to the method for the 17 mer probe described in part A.

400,000 recombinants are screened by the procedure described above in Part A. One duplicate positive is plaque purified and the DNA is isolated from a plate lysate of the recombinant bacteriophage designated lambda bP-21. Bacteriophage bP-21 was deposited with the ATCC under accession number ATCC 40310 on March 6, 1987. The bP-21 clone encodes the bovine growth factor designated bBMP-2.

The oligonucleotide hybridizing region of this bBMP-2 clone is localized to an approximately 1.2 kb Sac I restriction fragment which is subcloned into M13 and sequenced by standard techniques. The partial DNA sequence and derived amino acid sequence of this Sac I fragment and the contiguous Hind III-Sac I restriction fragment of bP-21 are shown below in Table III. The bBMP-2 peptide sequence from this clone is 129 amino acids in length and is encoded by the DNA sequence from nucleotide #1 through nucleotide #387. The amino acid sequence corresponding to the tryptic fragment isolated from the bovine bone 28 to 30kd material is underlined in Table III. The underlined portion of the sequence corresponds to tryptic Fragment 3 above from which the oligonucleotide probes for bBMP-2 are designed. The predicted amino acid sequence indicates that tryptic Fragment 3 is preceded by a basic

residue (K) as expected considering the specificity of trypsin. The arginine residue encoded by the CGT triplet is presumed to be the carboxy-terminus of the protein based on the presence of a stop codon (TAG) adjacent to it.

TABLE III

(1)                    15                    30                    45  
 GGC CAC GAT GGG AAA GGA CAC CCT CTC CAC AGA AGA GAA AAG CGG  
 G H D G K G H P L H R R E K R

                      60                    75                    90  
 CAA GCA AAA CAC AAA CAG CGG AAA CGC CTC AAG TCC AGC TGT AAG  
 Q A K H K Q R K R L K S S C K

                      105                    120                    135  
 AGA CAC CCT TTA TAT GTG GAC TTC AGT GAT GTG GGG TGG AAT GAC  
 R H P L Y V D F S D V G W N D

                      150                    165                    180  
 TGG ATC GTT GCA CCG CCG GGG TAT CAT GCC TTT TAC TGC CAT GGG  
 W I V A P P G Y H A F Y C H G

                      195                    210                    225  
 GAG TGC CCT TTT CCC CTG GCC GAT CAC CTT AAC TCC ACG AAT CAT  
 E C P F P L A D H L N S T N H

                      240                    255                    270  
 GCC ATT CTC CAA ACT CTG GTC AAC TCA GTT AAC TCT AAG ATT CCC  
 A I V Q T L V N S V N S K I P

                      385                    300                    315  
 AAG GCA TGC TGT GTC CCA ACA GAG CTC AGC GCC ATC TCC ATG CTG  
 K A C C V P T E L S A I S M L

                      330                    345                    360  
 TAC CTT GAT GAG AAT GAG AAG GTG GTA TTA AAG AAC TAT CAG GAC  
 Y L D E N E K V V L K N Y Q D

                      375                    (129)                    397                    407  
 ATG GTT GTC GAG GGT TGT GGG TGT CGT TAGCACAGCA AAATAAAATA  
M V V E G C G C R

                      417                    427                    437                    447                    457  
 TAAATATATA TATATATATA TTAGAAAAAC AGCAAAAAAA TCAAGTTGAC

                      467                    477                    487                    497                    507  
 ACTTTAATAT TTCCCAATGA AGACTTTATT TATGGAATGG AATGGAGAAA

                      517                    527                    537                    547                    557  
 AAGAAAAACA CAGCTATTTT GAAAACATAA TTTATATCTA CCGAAAAGAA

                      567                    577                    587  
 GTTGGGAAAAA CAAATATTTT AATCAGAGAA TTATT

C. bBMP-3

Probes consisting of pools of oligonucleotides are designed on the basis of the amino acid sequences of the tryptic Fragments 9 (Probe #3), 10 (Probe #2), and 11 (Probe #1), and synthesized on an automated DNA synthesizer.

Probe #1: A C N G T C A T [A/G] T T N G G [A/G] T A

Probe #2: C A [A/G] T A [A/G] T A N G C [A/G] T C [A/G] A A

Probe #3: T G [A/G/T] A T N G T N G C [A/G] T G [A/G] T T

A recombinant bovine genomic library constructed in EMBL3 is screened by the TMAC hybridization procedure detailed above in part A. 400,000 recombinants are screened in duplicate with Probe #1 which has been labeled with  $^{32}P$ . All recombinants which hybridized to this probe are replated for secondaries. Triplicate nitrocellulose replicas are made of the secondary plates, and amplified as described. The three sets of filters are hybridized to Probes #1, #2 and #3, again under TMAC conditions. One clone, lambda bP-819, hybridizes to all three probes and is plaque purified and DNA is isolated from a plate lysate. Bacteriophage lambda bP-819 was deposited with the ATCC on June 16, 1987 under accession number 40344. This bP-819 clone encodes the bovine bone growth factor designated bBMP-3.

The region of bP-819 which hybridizes to Probe #2 is localized and sequenced. The partial DNA and derived amino acid sequences of this region are shown in Table IVA. The amino acid sequences corresponding to tryptic Fragments 10 and 12 are underlined. The first underlined sequence corresponds to Fragment 12 while the second corresponds to Fragment 10. This region of bP-819, therefore, which hybridizes to Probe #2 encodes at least 111 amino acids. This amino acid sequence is encoded by the DNA sequence from nucleotide #414 through #746.

TABLE IV. A.

383	393	403	413	(1)	428
GAGGAGGAAG CGGTCTAOGG GGGTCCCTCT GCCTCTGCAG			AAC AAT GAG CTT CCT GGG GCA		
				Asn Asn Glu Leu Pro Gly Ala	
443	458		473		488
GAA TAT CAG TAC AAG GAG GAT GAA GTC TGG GAG AGG AAG CCT TAC AAG ACT					
Glu Tyr Gln Tyr Lys Glu Asp Glu Val Trp Glu Glu Arg Lys Pro Tyr Lys Thr					
503		518		533	
CIT CAG ACT CAG CCC CCT GAT AAG AGT AAG AAC AAA AAG AAA CAG AGG AAG GCA					
Leu Gln Thr Gln Pro Pro Asp Lys Ser Lys Asn Lys Lys Gln Arg Lys Gly					
548	563		578		593
CCT CAG CAG AAG AGT CAG ACG CTC CAG TTT GAT GAA CAG ACC CTG AAG AAG GCA					
Pro Gln Gln Lys Ser Gln Thr Leu Gln Phe Asp Glu Gln Thr Leu Lys Lys Ala					
608		623		638	
AGA AGA AAG CAA TGG ATT GAA CCC CGG AAT TGT GCC AGA CGG TAC CIT AAA GTG					
Arg Arg Lys Gln Trp Ile Glu Pro Arg Asn Cys Ala Arg Arg Tyr Leu Lys Val					
653	668		683		698
GAC TTC GCA GAT ATT CGC TGG AGC GAA TGG ATT ATT TCC CCC AAG TCC TTC GAT					
<u>Asp Phe Ala Asp Ile Gly Trp Ser Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp</u>					
713	728		743 (111)		756
GCC TAT TAC TGC TCC GGA CGG TGC CAG TTC CCC ATG CCA AAG GTAGCCATIG					
Ala Tyr Tyr Cys Ser Gly Ala Cys Gln Phe Pro MET Pro Lys					
766	776	786			
TTTTTGTCG TGTCTCTTCCC ATTTCCTATAG					

The region of bP-819 which hybridizes to Probe #1 and #3 is localized and sequenced. The partial DNA and derived amino acid sequences of this region are shown in Table IVB. The amino acid sequences corresponding to tryptic Fragments 9 and 11 are underlined. The first underlined sequence corresponds to Fragment 9 while the second underlined sequence corresponds to Fragment 11. The peptide sequence of this region of bP-819 which hybridizes to Probe #1 and #3 is 64 amino acids in length encoded by nucleotide #305 through #493 of Table IVB. The arginine residue encoded by the AGA triplet is presumed to be the carboxy-terminus of the protein based on the presence of a stop codon (TAA) adjacent to it. The nucleic acid sequence preceding the couplet TC (positions 305-306) is presumed to be an intron (non-coding sequence) based on the presence of a consensus acceptor sequence (i.e. a pyrimidine-rich stretch, TTCTCCCTTTCGTTCCCT, followed by AG) and the presence of a stop rather than a basic residue in the appropriate position of the derived amino acid sequence.

bBMP-3 is therefore characterized by the DNA and amino acid sequence of Table IV A and Table IV B. The peptide sequence of this clone is 175 amino acids in length and is encoded by the DNA sequence from nucleotide #414 through nucleotide #746 of Table IV A and nucleotide #305 through nucleotide #493 of Table IV B.

TABLE IV. B.

284	294	304 (112)	319
CTAACCGTGT TTCCTCCCTT TGGTCCCTAG TCT TTG AAG CCA TCA AAT CAC GCT ACC			
<u>Ser Leu Lys Pro Ser Asn His Ala Thr</u>			
334	349	364	379
ATC CAG AGT ATA GTG ACA GCT GTG GGG GTC GTC CCT GGA ATC CCC GAG CCT TGC			
<u>Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile Pro Glu Pro Cys</u>			
394	409	424	439
TGT GTG CCA GAA AAG ATG TCC TCA CTC AGC ATC TTA TTC TTT GAT GAA AAC AAG			
<u>Cys Val Pro Glu Lys MET Ser Ser Leu Ser Ile Leu Phe Phe Asp Glu Asn Lys</u>			
454	469	484	(175)
AAT GTG GTA CTT AAA GTA TAT CCA AAC ATG ACA GTA GAG TCT TGT GCT TGC AGA			
<u>Asn Val Val Leu Lys Val Tyr Pro Asn MET Thr Val Glu Ser Cys Ala Cys Arg</u>			
503	513	523	533
TAACCGTGT AAGAACTCAT CTGGATGCCT AACCTCAATCG			

EXAMPLE VHuman Bone Inductive FactorsA. hBMP-1

Because the bovine and human bone growth factor genes are presumed to be significantly homologous, the bovine bBMP-1 DNA sequence of Table II (or portions thereof) is used as a probe to screen a human genomic library. The 800bp EcoRI fragment of the bovine genomic clone is labeled with  $^{32}\text{P}$  by nick-translation. A human genomic library (Toole et al., supra) is plated on 20 plates at 40,000 recombinants per plate. Duplicate nitrocellulose filter replicas are made of each plate and hybridized to the nick-translated probe in 5 X SSC, 5 X Denhardt's, 100ug/ml denatured salmon sperm DNA, 0.1% SDS (the standard hybridization solution) at 50 degrees centigrade for approximately 14 hours. The filters are then washed in 1 X SSC, 0.1% SDS at 50 degrees centigrade and subjected to autoradiography. Five duplicate positives are isolated and plaque purified. DNA is obtained from a plate lysate of one of these recombinant bacteriophage, designated LP-H1. LP-H1 was deposited with the ATCC on March 6, 1987 under accession number 40311. This clone encodes at least a portion of the human genomic bone growth factor called hBMP-1. The hybridizing region of LP-H1 is localized to a 2.5kb XbaI/HindIII restriction fragment.

The partial DNA sequence and derived amino acid sequence of lambda LP-H1 are shown below in Table V. The peptide sequence from this clone is 37 amino acids in length and is encoded by the DNA sequence from nucleotide #3440 through nucleotide #3550. The coding sequence of Table V is flanked by approximately 28 nucleotides (a presumptive 5' noncoding sequence) as well as approximately 19 nucleotides (a presumptive 3' noncoding sequence). A comparison of the bBMP-1 sequence of Table II with the hBMP-1 genomic sequence of Table V indicates the significant homology between the two.

Because the size of coding regions and the positions

of noncoding regions is generally conserved in homologous genes of different species, the locations of the coding and noncoding regions of the bone inductive factor genes may be identified. Regions of homology between the two species' genes, flanked by RNA processing signals at homologous sites, indicate a coding region.

TABLE V

3419	3429	3439	(1)	3454
CAGCOCTGGC	TTCCTTCCTTTT	CICCTTAGCT	GOC TTT CTT CGG GAC ATT GOC CTG GAC	
			Ala Phe Leu Gly Asp Ile Ala Leu Asp	
3469	3484		3499	3514
GAA GAG GAC	CIG AGG GOC	TTC CAG	GTA CAG CAG GCT GIG GAT CTC AGA CGG CAC	
Glu Glu Asp	Leu Arg Ala	Phe Gln Val	Gln Gln Ala Val Asp	Leu Arg Arg His
3529	3544	(37)	3560	3570
ACA GCT CGT AAG	TCC TCC ATC AAA GCT	GCA GGTAAGCGG	GTGCCAATGG	
Thr Ala Arg Lys Ser	Ser Ile Lys Ala	Ala Ala		

A probe specific for the human coding sequence given in Table V is used to identify a human cell line or tissue which synthesizes bone inductive factor. The probe is made according to the following method. Two oligonucleotides having the following sequences:

- (a) GGGATTCTGCCTTTCTGGGGACATTGCCCTGGACGAAGAGGACCTGAG  
(b) CGGGATCCGTCTGAGATCCACAGCCTGCTGTACCTGGAAGGCCCTCAGG

are synthesized on an automated synthesizer, annealed, extended using the Klenow fragment of *E. coli* DNA polymerase I, digested with the restriction enzymes Eco RI and Bam HI, and inserted into an M13 vector. A single-stranded <sup>32</sup>P-labeled probe is then from template preparation of this subclone by standard techniques. Polyadenylated RNAs from various cell and tissue sources are electrophoresed on formaldehyde-agarose gels and transferred to nitrocellulose by the method of Toole et al., supra. The probe is then hybridized to the nitrocellulose blot in 50% formamide, 5 X SSC, 0.1% SDS, 40 mM sodium phosphate pH 6.5, 100 ug/ml denatured salmon sperm DNA, and 5 mM vanadyl ribonucleosides at 42° C overnight and washed at 65° C in 0.2 X SSC, 0.1% SDS. Following autoradiography, the lane containing RNA from the human osteosarcoma cell line U-2 OS contains hybridizing bands corresponding to RNA species of approximately 4.3 and 3.0 kb.

cDNA is synthesized from U-2 OS polyadenylated RNA and cloned into lambda gt10 by established techniques (Toole et al., supra). 20,000 recombinants from this library are plated on each of 50 plates. Duplicate nitrocellulose replicas are made of the plates. The above described oligonucleotides are kinased with <sup>32</sup>P-gamma-ATP and hybridized to the two sets of replicas at 55° centigrade in standard hybridization solution overnight. The filters are then washed in 1 X SSC, 0.1% SDS at 55° centigrade and subjected to autoradiography. One duplicate positive, designated lambda U2OS-1, is plaque purified. Lambda U2OS-1 was deposited with the ATCC on June 16, 1987 under accession number 40343.

The entire nucleotide sequence and derived amino acid sequence of the insert of lambda U2OS-1 is given in Table VI. This cDNA clone encodes a Met followed by a hydrophobic leader sequence characteristic of a secreted protein, and contains a stop codon at nucleotide positons 2226 - 2228. This clone contains an open reading frame of 2190bp, encoding a protein of 730 amino acids with a molecular weight of 83kd based on this amino acid sequence. The clone contains sequence identical to the coding region given in Table V. This protein is contemplated to represent a primary translation product which is cleaved upon secretion to produce the hBMP-1 protein. This clone is therefore a cDNA for hBMP-1 corresponding to human gene fragment contained in the genomic hBMP-1 sequence lambda LP-H1. It is noted that amino acids #550 to #590 of BMP-1 are homologous to epidermal growth factor and the "growth factor" domains of Protein C, Factor X and Factor IX.

TABLE VI

10	20	30	(1)	50
CTAGAGGCG CTTCCCTGC CGCGCGCG CGAC ATG CCC GGC GIG GGC CGC CTG CG MET Pro Gly Val Ala Arg Leu Pro				
65	80	95		110
CTG CTG CTC GGG CTG CTG CTG CGT CCC GGC CGG CGG CTG GAC TTG GGC Leu Leu Leu Gly Leu Leu Leu Pro Arg Pro Gly Arg Pro Leu Asp Leu Ala				
125	140		155	
GAC TAC ACC TAT GAC CTG CGG GAG GAG GAC GAC TOG GAG CCC CTC AAC TAC AAA Asp Tyr Thr Tyr Asp Leu Ala Glu Glu Asp Asp Ser Glu Pro Leu Asn Tyr Lys				
170	185	200		215
GAC CCC TGC AAG GCG GCT GCC TTT CCT GGG GAC ATT GCC CTG GAC GAA GAG GAC Asp Pro Cys Lys Ala Ala Ala Phe Leu Gly Asp Ile Ala Leu Asp Glu Glu Asp				
230	245	260		275
CTG AGG GCC TTC CAG GTC CAG CAG GCT GIG GAT CTC AGA CGG CAC ACA GCT CGT Leu Arg Ala Phe Gln Val Gln Ala Val Asp Leu Arg Arg His Thr Ala Arg				
290	305	320		
AAG TCC TCC ATC AAA GCT GCA GTF CCA GGA AAC ACT TCT ACC CCC AGC TGC CAG Lys Ser Ser Ile Lys Ala Ala Val Pro Gly Asn Thr Ser Thr Pro Ser Cys Gln				
335	350	365		380
AGC ACC AAC GGG CAG CCT CAG AGG GGA GCC TGT GGG AGA TGG AGA GGT AGA TCC Ser Thr Asn Gly Gln Pro Gln Arg Gly Ala Cys Gly Arg Trp Arg Gly Arg Ser				
395	410	425		
GCT AGC CGG CGG GCG GCG ACG TCC CGA CCA GAG CGT GTG TGG CCC GAT GGG GTC Arg Ser Arg Arg Ala Ala Thr Ser Arg Pro Glu Arg Val Trp Pro Asp Gly Val				
440	455	470		485
ATC CCC TTT GTC ATT GGG GGA AAC TTC ACT GGT AGC CAG AGG GCA GTC TTC CGG Ile Pro Phe Val Ile Gly Gly Asn Phe Thr Gly Ser Gln Arg Ala Val Phe Arg				
500	515	530		545
CAG GGC ATG AGC CAC TGG GAG AAG CAC ACC TGT GTC ACC TTC CTG GAG CGC ACT Gln Ala MET Arg His Trp Glu Lys His Thr Cys Val Thr Phe Leu Glu Arg Thr				
560	575	590		
GAC GAG GAC AGC TAT ATT GIG TTC ACC TAT CGA CCT TGC GGG TGC TGC TCC TAC Asp Glu Asp Ser Tyr Ile Val Phe Thr Tyr Arg Pro Cys Gly Cys Ser Tyr				
605	620	635		650
GTC GGT CGC CGC GGC GGG GGC CCC CAG GCC ATC TCC ATC GGC AAG AAC TGT GAC Val Gly Arg Arg Gly Gly Pro Gln Ala Ile Ser Ile Gly Lys Asn Cys Asp				

665                    680                    695  
 AAG TTC GGC ATT GTG GTC CAC GAG CTG GGC CAC GTC GTC GGC TTC TGG CAC GAA  
 Lys Phe Gly Ile Val Val His Glu Leu Gly His Val Val Gly Phe Trp His Glu  
  
 710                    725                    740                    755  
 CAC ACT CGG CCA GAC CGG GAC OGC CAC GTT TCC ATC GTT CGT GAG AAC ATC CAG  
 His Thr Arg Pro Asp Arg Asp His Val Ser Ile Val Arg Glu Asn Ile Gln  
  
 770                    785                    800                    815  
 CCA GGG CAG GAG TAT AAC TTC CTG AAG ATG GAG CCT CAG GAG GTG GAG TCC CTG  
 Pro Gly Gln Glu Tyr Asn Phe Leu Lys MET Glu Pro Gln Glu Val Glu Ser Leu  
  
 830                    845                    860  
 GGG GAG ACC TAT GAC TTC GAC AGC ATC ATG CAT TAC GCT CGG AAC ACA TTC TCC  
 Gly Glu Thr Tyr Asp Phe Asp Ser Ile MET His Tyr Ala Arg Asn Thr Phe Ser  
  
 875                    890                    905                    920  
 AGG GCC ATC TTC CTG GAT ACC ATT GTC CCC AAG TAT GAG GTG AAC GGG GTG AAA  
 Arg Gly Ile Phe Leu Asp Thr Ile Val Pro Lys Tyr Glu Val Asn Gly Val Lys  
  
 935                    950                    965  
 CCT CCC ATT GGC CAA AGG ACA CGG CTC AGC AAG GGG GAC ATT GCC CAA GCC CGC  
 Pro Pro Ile Gly Gln Arg Thr Arg Leu Ser Lys Gly Asp Ile Ala Gln Ala Arg  
  
 980                    995                    1010                    1025  
 AAG CTT TAC AAG TGC CCA GCC TGT GGA GAG ACC CTG CAA GAC AGC ACA GGC AAC  
 Lys Leu Tyr Lys Cys Pro Ala Cys Gly Glu Thr Leu Gln Asp Ser Thr Gly Asn  
  
 1040                    1055                    1070                    1085  
 TTC TCC TCC CCT GAA TAC CCC AAT GGC TAC TCT GCT CAC ATG CAC TGC GTG TGG  
 Phe Ser Ser Pro Glu Tyr Pro Asn Gly Tyr Ser Ala His MET His Cys Val Trp  
  
 1100                    1115                    1130  
 CGC ATC TCT GTC ACA CCC GGG GAG AAG ATC ATC CTG AAC TTC ACG TOC CTG GAC  
 Arg Ile Ser Val Thr Pro Gly Glu Ile Ile Leu Asn Phe Thr Ser Leu Asp  
  
 1145                    1160                    1175                    1190  
 CTG TAC CGC AGC CGC CTC TGC TGG TAC GAC TAT GTG GAG GTC CGA GAT GGC TTC  
 Leu Tyr Arg Ser Arg Leu Cys Trp Tyr Asp Tyr Val Glu Val Arg Asp Gly Phe  
  
 1205                    1220                    1235  
 TGG AGG AAG GCG CCC CTC CGA GGC CGC TTC TGC GGG TOC AAA CTC CCT GAG CCT  
 Trp Arg Lys Ala Pro Leu Arg Gly Arg Phe Cys Gly Ser Lys Leu Pro Glu Pro  
  
 1250                    1265                    1280                    1295  
 ATC GTC TOC ACT GAC AGC CGC CTC TGG GTT GAA TTC CGC AGC AGC AGC AAT TGG  
 Ile Val Ser Thr Asp Ser Arg Leu Trp Val Glu Phe Arg Ser Ser Asn Trp  
  
 1310                    1325                    1340                    1355  
 GTT GGA AAG GGC TTC TTT GCA GTC TAC GAA GCC ATC TGC GGG GGT GAT GTG AAA  
 Val Gly Lys Gly Phe Phe Ala Val Tyr Glu Ala Ile Cys Gly Gly Asp Val Lys

1370                    1385                    1400  
 AAG GAC TAT GCC CAC ATT CAA TCG CCC AAC TAC CCA GAC GAT TAC CGG CCC AGC  
 Lys Asp Tyr Gly His Ile Gln Ser Pro Asn Tyr Pro Asp Asp Tyr Arg Pro Ser  
  
 1415                    1430                    1445                    1460  
 AAA GTC TGC ATC TGG CGG ATC CAG GTG TCT GAG GGC TTC CAC GTG GGC CTC ACA  
 Lys Val Cys Ile Trp Arg Ile Gln Val Ser Glu Gly Phe His Val Gly Leu Thr  
  
 1475                    1490                    1505  
 TTC CAG TCC TTT GAG ATT GAG CCC CAC GAC AGC TGT GCC TAC GAC TAT CTG GAG  
 Phe Gln Ser Phe Glu Ile Glu Arg His Asp Ser Cys Ala Tyr Asp Tyr Leu Glu  
  
 1520                    1535                    1550                    1565  
 GIG CGC GAC GGG CAC AGT GAG AGC AGC ACC CTC ATC GGG CGC TAC TGT GGC TAT  
 Val Arg Asp Gly His Ser Glu Ser Ser Thr Leu Ile Gly Arg Tyr Cys Gly Tyr  
  
 1580                    1595                    1610                    1625  
 GAG AAG CCT GAT GAC ATC AAG AGC ACG TCC AGC CGC CTC TGG CTC AAG TTC GTC  
 Glu Lys Pro Asp Asp Ile Lys Ser Thr Ser Ser Arg Leu Trp Leu Lys Phe Val  
  
 1640                    1655                    1670  
 TCT GAC GGG TCC ATT AAC AAA GCG GGC TTT GCC GTC AAC TTT TTC AAA GAG GTG  
 Ser Asp Gly Ser Ile Asn Lys Ala Gly Phe Ala Val Asn Phe Phe Lys Glu Val  
  
 1685                    1700                    1715                    1730  
 GAC GAG TGC TCT CGG CCC AAC CGC GGG GGC TGT GAG CAG CGG TGC CTC AAC ACC  
 Asp Glu Cys Ser Arg Pro Asn Arg Gly Gly Cys Glu Gln Arg Cys Leu Asn Thr  
  
 1745                    1760                    1775  
 CTG CGC AGC TAC AAG TGC AGC TGT GAC CCC GGG TAC GAG CTG GCC CCA GAC AAG  
 Leu Gly Ser Tyr Lys Cys Ser Cys Asp Pro Gly Tyr Glu Leu Ala Pro Asp Lys  
  
 1790                    1805                    1820                    1835  
 CGC CGC TGT GAG CCT GCT TGT GGC GGA TTC CTC ACC AAG CTC AAC GGC TCC ATC  
 Arg Arg Cys Glu Ala Ala Cys Gly Gly Phe Leu Thr Lys Leu Asn Gly Ser Ile  
  
 1850                    1865                    1880                    1895  
 ACC AGC CGG GCC TGG CCC AAG GAG TAC CCC CCC AAC AAG AAC TGC ATC TGG CAG  
 Thr Ser Pro Gly Trp Pro Lys Glu Tyr Pro Pro Asn Lys Asn Cys Ile Trp Gln  
  
 1910                    1925                    1940  
 CTG GTG GCC CCC ACC CAG TAC CGC ATC TCC CTG CAG TTT GAC TTC TTT GAG ACA  
 Leu Val Ala Pro Thr Gln Tyr Arg Ile Ser Leu Gln Phe Asp Phe Glu Thr  
  
 1955                    1970                    1985                    2000  
 GAG GGC AAT GAT GTG TGC AAG TAC GAC TTC GTG GAG GTG CGC AGT GCA CTC ACA  
 Glu Gly Asn Asp Val Cys Lys Tyr Asp Phe Val Glu Val Arg Ser Gly Leu Thr  
  
 2015                    2030                    2045  
 GCT GAC TCC AAG CTG CAT GGC AAG TTC TGT GGT TCT GAG AAG CCC GAG GTC ATC  
 Ala Asp Ser Lys Leu His Gly Lys Phe Cys Gly Ser Glu Lys Pro Glu Val Ile

2060                  2075                  2090                  2105  
ACC TCC CAG TAC AAC AAC ATG CGC GTG GAG TTC AAG TCC GAC AAC ACC GTG TCC  
Thr Ser Gln Tyr Asn Asn MET Arg Val Glu Phe Lys Ser Asp Asn Thr Val Ser

2120                  2135                  2150                  2165  
AAA AAG GGC TTC AAG GCC CAC TTC TTC TCA GAA AAG AGG CCA GCT CTG CAG CCC  
Lys Lys Gly Phe Lys Ala His Phe Phe Ser Glu Lys Arg Pro Ala Leu Gln Pro

2180                  2195                  2210  
CCT CCG GGA CGC CCC CAC CAG CTC AAA TTC CGA GTG CAG AAA AGA AAC CCG ACC  
Pro Arg Gly Arg Pro His Gln Leu Lys Phe Arg Val Gln Lys Arg Asn Arg Thr  
(730)

2225                  2235                  2245                  2255                  2265                  2275                  2285  
CCC CAG TGAGGCGTGC CAGGCGCTOC GGACCCCTTG TTACTCAGGA ACCTCACCTT GGACGGAAATG  
Pro Gln

2295                  2305                  2315                  2325                  2335                  2345                  2355  
GGATGGGGGC TTGGGTGCCC ACCAACCCCC CACCTCCACT CTGOCATTCC GGOCCACCTC CCTCTGGCOG

2365                  2375                  2385                  2395                  2405                  2415                  2425  
GACAGAACCTG GTGCTCTCTT CTCCCCACTG TGCGGCGTGG OGGACGGGG ACCCTTCCCCC GTGCGCTACC

2435                  2445                  2455                  2465                  2475                  2485                  2495  
CCCTCCCACTT TTGATGGGTG CTGAGACATT TCCCTGTTG AAGTAAAGCA GGGACCCCTG CGTCTGGCT

CTAGA

B. hBMP-2: Class I and II

The HindIII-SacI bovine genomic bBMP-2 fragment described in Example IV B. is subcloned into an M13 vector. A  $^{32}\text{P}$ -labeled single-stranded DNA probe is made from a template preparation of this subclone. This probe is used to screen polyadenylated RNAs from various cell and tissue sources as described above in part A. A hybridizing band corresponding to an mRNA species of approximately 3.8 kb is detected in the lane containing RNA from the human cell line U-2 OS. The HindIII-SacI fragment is labeled with  $^{32}\text{P}$  by nick translation and used to screen the nitrocellulose filter replicas of the above-described U-2 OS cDNA library by hybridization in standard hybridization buffer at 65° overnight followed by washing in 1 X SSC, 0.1% SDS at 65°. Twelve duplicate positive clones are picked and replated for secondaries. Duplicate nitrocellulose replicas are made of the secondary plates and both sets hybridized to the bovine genomic probe as the primary screening was performed. One set of filters is then washed in 1 X SSC, 0.1% SDS; the other in 0.1 X SSC, 0.1% SDS at 65°.

Two classes of hBMP-2 cDNA clones are evident based on strong (4 recombinants) or weak (7 recombinants) hybridization signals under the more stringent washing conditions (0.1 X SSC, 0.1% SDS). All 11 recombinant bacteriophage are plaque purified, small scale DNA preparations made from plate lysates of each, and the inserts subcloned into pSP65 and into M13 for sequence analysis. Sequence analysis of the strongly hybridizing clones designated hBMP-2 Class I (also known as BMP-2) indicates that they have extensive sequence homology with the sequence given in Table III. These clones are therefore cDNA encoding the human equivalent of the protein encoded by the bBMP-2 gene whose partial sequence is given in Table III. Sequence analysis of the weakly hybridizing recombinants designated hBMP-2 Class II (also known as BMP-4) indicates that they are also quite homologous

with the sequence given in Table III at the 3' end of their coding regions, but less so in the more 5' regions. Thus they encode a human protein of similar, though not identical, structure to that above.

Full length hBMP-2 Class I cDNA clones are obtained in the following manner. The 1.5 kb insert of one of the Class II subclones (II-10-1) is isolated and radioactively labeled by nick-translation. One set of the nitrocellulose replicas of the U-2 OS cDNA library screened above (50 filters, corresponding to 1,000,000 recombinant bacteriophage) are rehybridized with this probe under stringent conditions (hybridization at 65° in standard hybridization buffer; washing at 65° in 0.2 X SSC, 0.1% SDS). All recombinants which hybridize to the bovine genomic probe which do not hybridize to the Class II probe are picked and plaque purified (10 recombinants). Plate stocks are made and small scale bacteriophage DNA preparations made. After subcloning into M13, sequence analysis indicates that 4 of these represent clones which overlap the original Class I clone. One of these, Lambda U2OS-39, contains an approximately 1.5 kb insert and was deposited with the ATCC on June 16, 1987 under accession number 40345. The partial DNA sequence (compiled from Lambda U2OS-39 and several other hBMP-2 Class I cDNA recombinants) and derived amino acid sequence are shown below in Table VII. Lambda U2OS-39 is expected to contain all of the nucleotide sequence necessary to encode the entire human counterpart of the protein BMP-2 Class II encoded by the bovine gene segment whose partial sequence is presented in Table III. This human cDNA hBMP-2 Class II contains an open reading frame of 1188 bp, encoding a protein of 396 amino acids. This protein of 396 amino acids has a molecular weight of 45kd based on this amino acid sequence. It is contemplated that this sequence represents the primary translation product. The protein is preceded by a 5' untranslated region of 342 bp with stop codons in all frames.

40

The 13 bp region preceding this 5' untranslated region represents a linker used in the cDNA cloning procedure.

TABLE VII

10	20	30	40	50	60	70
GTOGACTCTA GAGTGTGTGT CAGCACCTGG CTGGGGACTT CTGAACTTG CAGGGAGAAT AACATTGCGCA						
80	90	100	110	120	130	140
CCOCACCTTG CGCGGGTGCC TTGCCCCAG CGGAGCTTCG TTGCGCATCT CGAGCGGCA CGCGCCCTCC						
150	160	170	180	190	200	210
ACTCCTGGC CTGCGGGAC ACTGAGAOGC TGTCCCCAGC GTGAAAAGAG AGACCTGGGGC GCGGGCACCC						
220	230	240	250	260	270	280
GGGAGAAGGA GGAGGCAAAG AAAAGGAAOG GACATTGGT CCTTGGCGCA GGTCCCTTGA CCAGAGTTT						
290	300	310	320	330	340	350
TCCATGTGGA CGCTCTTTCA ATGGAOGTGT CGCGCGTGC TTCTTAAOG GACCTGGGT TOCTAAAGGT						
(1)	370		385		400	
CGACC ATG GTG GCC EGG ACC CGC TGT CTT CTA CGC TTG CTG CTT CCC CAG GTC MET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Val						
415		430		445		
CTC CTG GGC GGC CGT GCT GGC CTC GTC CGT GAG CTG GGC CGC AGG AAG TTC GCG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe Ala						
460	475		490		505	
GCG GCG TCG TCG CGC CGC CGC TCA TCC CAG CGC TCT GAC GAG GTC CTG AGC GAG Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val Leu Ser Glu						
520		535		550		565
TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG AGA CGC ACC CGC ACC AGC Phe Glu Leu Arg Leu Leu Ser MET Phe Gly Leu Lys Gln Arg Pro Thr Pro Ser						
580		595		610		
AGG GAC GGC GTG GTG CGC CGC TAC ATG CTA GAC CTG TAT CGC AGG CAC TCG GGT Arg Asp Ala Val Val Pro Pro Tyr MET Leu Asp Leu Tyr Arg Arg His Ser Gly						
625		640		655		670
CAG CGG CGC TCA CGC CGC CGA GAC CAC CGG TTG GAG AGG GCA CGC AGC CGA CGC Gln Pro Gly Ser Pro Ala Pro Asp His Arg Leu Glu Arg Ala Ala Ser Arg Ala						
685		700		715		
AAC ACT GTG CGC AGC TTC CAC CAT GAA GAA TCT TTG GAA GAA CTA CGA GAA ACG Asn Thr Val Arg Ser Phe His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr						

730                745                760                775  
 AGT GGG AAA ACA ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG  
 Ser Gly Lys Thr Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu  
  
 790                805                820                835  
 CAG TTT ATC ACC TCA GCA GAG CCT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT  
 Glu Phe Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln MET Gln Asp Ala  
  
 850                865                880  
 TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA AAA  
 Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile Lys  
  
 895                910                925                940  
 CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGT CCT TTG GAC ACC AGG TTG  
 Pro Ala Thr Ala Asn Ser Phe Pro Val Thr Ser Leu Leu Asp Thr Arg Leu  
  
 955                970                985  
 GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT GTC ACC CCC GCT GTG ATG  
 Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp Val Thr Pro Ala Val MET  
  
 1000              1015              1030              1045  
 CCG TGG ACT GCA CAG GGA CAC GCC AAC CAT GGA TTC GTG GTG GAA GTG GCC CAC  
 Arg Trp Thr Ala Gln Gly His Ala Asn His Gly Phe Val Val Glu Val Ala His  
  
 1060              1075              1090              1105  
 TTG GAG GAG AAA CAA GGT GTC TCC AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG  
 Leu Glu Glu Lys Gln Gly Val Ser Lys Arg His Val Arg Ile Ser Arg Ser Leu  
  
 1120              1135              1150  
 CAC CAA GAT GAA CAC AGC TGG TCA CAG ATA AGG CCA TTG CTC GTC ACT TTT GGC  
 His Gln Asp Glu His Ser Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly  
  
 1165              1180              1195              1210  
 CAT GAT GGA AAA GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC  
 His Asp Gly Lys His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His  
  
 1225              1240              1255  
 AAA CAG CGG AAA CGC CCT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC  
 Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp  
  
 1270              1285              1300              1315  
 TTC AGT GAC GTG CGG TGG AAT GAC TGG ATT GTG GCT CCC CGG GGG TAT CAC GCC  
 Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala  
  
 1330              1345              1360              1375  
 TTT TAC TGC CAC CGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG AAC TCC ACT  
 Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr  
  
 1390              1405              1420  
 AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC TCT AAG ATT CCT AAG  
 Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys

43

1435            1450            1465            1480  
GCA TGC TGT GTC CGC ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG  
Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu

1495            1510            1525  
AAT GAA AAG GTT GTA TTA AAG AAC TAT CAG GAC ATG GTT GTG GAG GGT TGT GGG  
Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly

1540(396)      1553      1563      1573      1583      1593      1603  
TGT CGC TAGTACAGCA AAATTAAATA CATAAAATATA TATATATATA TATATTTAG AAAAAAGAAA  
Cys=Arg:

AAAAA

Full-length hBMP-2 Class II human cDNA clones are obtained in the following manner. The 200 bp EcoRI-SacI fragment from the 5' end of the Class II recombinant II-10-1 is isolated from its plasmid subclone, labeled by nick-translation, and hybridized to a set of duplicate nitrocellulose replicas of the U-2 OS cDNA library (25 filters/set; representing 500,000 recombinants). Hybridization and washing are performed under stringent conditions as described above. 16 duplicate positives are picked and replated for secondaries. Nitrocellulose filter replicas of the secondary plates are made and hybridized to an oligonucleotide which was synthesized to correspond to the sequence of II-10-1 and is of the following sequence:

CGGGCGCTCAGGATACTCAAGACCAGTGCTG

Hybridization is in standard hybridization buffer AT 50° C with washing at 50° in 1 X SSC, 0.1% SDS. 14 recombinant bacteriophage which hybridize to this oligonucleotide are plaque purified. Plate stocks are made and small scale bacteriophage DNA preparations made. After subcloning 3 of these into M13, sequence analysis indicates that they represent clones which overlap the original Class II clone. One of these, lambda U2OS-3, was deposited with the ATCC under accession number 40342 on June 16, 1987. U2OS-3 contains an insert of approximately 1.8 kb. The partial DNA sequence and derived amino acid sequence of U2OS-3 are shown below in Table VIII. This clone is expected to contain all of the nucleotide sequence necessary to encode the entire human BMP-2 Class II protein. This cDNA contains an open reading frame of 1224 bp, encoding a protein of 408 amino acids, preceded by a 5' untranslated region of 394 bp with stop codons in all frames, and contains a 3' untranslated region of 308 bp following the in-frame stop codon. The 8 bp region preceding the 5' untranslated region represents a linker used in the cDNA cloning procedure. This protein of 408 amino acids has molecular weight of 47kd and is contemplated to represent the

45

primary translation product.

TABLE VIII

10	20	30	40	50	60	70
CCTCTAGAGGG CAGAGGACCA CGGAGGGAGG GAAGGAGGCGG GGAGCCCGGC COGGAAGCTA GGIGAGTGTC						
80	90	100	110	120	130	140
GCATCOGAGC TGAGGGAGC GAGCTGAGA CGCGCTGCT GCCTCGGCTG AGTATCTAGC TIGCTCOC						
150	160	170	180	190	200	210
CATGGGATTC CGGCTCAAGC TATCTGAGC CTGGAGGCG ACAGTCCCGG GCGCTCGGCC AGGTTCACIG						
220	230	240	250	260	270	280
CAACCGTTCA GAGGTTCCCA CGACCTTGCTG CTGGCGAGCC CGCTACCTGCA GGGACCTATG GAGCCATTCC						
290	300	310	320	330	340	350
GTAGTGCCAT CGCGAGCAAC GCACTGCTGC AGCTTCCCTG AGCCTTTCCA GCAAGTTTGT TCAAGATTTG						
360	370	380	390	400	(1)	
CTGTCAGAAA TCATGGACIG TTATTTATAG CCTTTTTTC TGTCAGACA CC ATG ATT CCT MET Ile Pro						
417	432	447	462			
GGT AAC CGA ATG CTG ATG GTC GTT TTA TTA TGC CAA GTC CTG CTA GGA GCC GCG Gly Asn Arg MET Leu Val Val Leu Cys Gln Val Leu Leu Gly Gly Ala						
477	492	507				
AGC CAT GCT AGT TIG ATA CCT GAG ACG GGG AAG AAA AAA GTC GCC GAG ATT CAG Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys Lys Val Ala Glu Ile Gln						
522	537	552	567			
GGC CAC GCG GGA GGA CGC CGC TCA GGG CAG AGC CAT GAG CTC CTG CGG GAC TTC Gly His Ala Gly Arg Arg Ser Gly Gln Ser His Glu Leu Leu Arg Asp Phe						
582	597	612	627			
GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGC CGC CGG CAG CCT AGC AAG Glu Ala Thr Leu Leu Gln MET Phe Gly Leu Arg Arg Pro Gln Pro Ser Lys						
642	657	672				
AGT GCC GTC ATT CGG GAC TAC ATG CGG CAT CTT TAC CGG CTT CAG TCT GGG GAG Ser Ala Val Ile Pro Asp Tyr MET Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu						
687	702	717	732			
GAG GAG GAA GAG CAG ATC CAC AGC ACT GGT CTT GAG TAT CCT GAG CGC CGG GCG Glu Glu Glu Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala						

47

747	762	777	
AGC CGG GCC AAC ACC GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC Ser Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile			
792	807	822	837
CCA GGG ACC AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile			
852	867	882	897
CCT GAG AAC GAG GTG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG CTG Pro Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln Val			
912	927	942	
GAC CAG GCC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT TAT GAG GTT Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile Tyr Glu Val			
957	972	987	1002
ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC ACA CGA CTA CTG GAC MET Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile Thr Arg Leu Leu Asp			
1017	1032	1047	
ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG GAA ACT TTT GAT GTG AGC CCT Thr Arg Leu Val His His Asn Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro			
1062	1077	1092	1107
GCG GTC CCT CGC TGG ACC CGG GAG AAG CAG CCA AAC TAT GGG CTA GCC ATT GAG Ala Val Leu Arg Trp Thr Arg Glu Lys Pro Asn Tyr Gly Leu Ala Ile Glu			
1122	1137	1152	1167
GTG ACT CAC CTC CAT CAG ACT CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC Val Thr His Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser			
1182	1197	1212	
CGA TCG TTA CCT CAA GGG AGT GGG AAT TGG GOC CAG CTC CGG CCC CTC CTG GTC Arg Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val			
1227	1242	1257	1272
ACC TTT GGC CAT GAT GGC CGG GGC CAT GOC TTG ACC CGA CGC CGG AGG GCC AAG Thr Phe Gly His Asp Gly Arg His Ala Leu Thr Arg Arg Arg Ala Lys			
1287	1302	1317	
CGT AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC CGG Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys Arg			
1332	1347	1362	1377
CGC CAC TCG CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC TGG ATT GTG Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val			
1392	1407	1422	1437
GCC CCA CCA GGC TAC CAG GOC TTC TAC TGC CAT GGG GAC TGC CGC TTT CCA CTG Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp Cys Pro Phe Pro Leu			

1452                    1467                    1482  
 GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT GIG CAG ACC CIG GTC AAT TCT  
 Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser  
  
 1497                    1512                    1527                    1542  
 GTC AAT TCC AGT ATC CCC AAA GCC TGT TGT GIG CCC ACT GAA CTG AGT GCC ATC  
 Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile  
  
 1557                    1572                    1587  
 TCC ATG CTG TAC CTG GAT GAG TAT GAT AAG GIG GTA CTG AAA AAT TAT CAG GAG  
 Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu  
  
 1602                    1617                    (408)                    1636                    1646                    1656  
 ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTOCTGAA GGATAGACAG  
 MET Val Val Glu Gly Cys Gly Cys Arg  
  
 1666                    1676                    1686                    1696                    1706                    1716                    1726  
 ATATACACAC CACACACACA CACCACATAC ACCACACACA CAOGTTCCA TCCACTCAAC CACACACTAC  
  
 1736                    1746                    1756                    1766                    1776                    1786                    1796  
 ACAGACTGCT TCCCTTATAGC TGGACTTTTA TTTAAAAAAA AAAAAAAA AATGGAAAAA ATCCCTAAAC  
  
 1806                    1816                    1826                    1836                    1846                    1856                    1866  
 ATTCAACCTTG ACCTTTATTA TGACTTTACG TGCAAATGTT TTGACCATAT TGATCATATA TTTTGACAAA  
  
 1876                    1886                    1896                    1906                    1916                    1926                    1936  
 ATATATTAT AACTAAGTAT TAAAAGAAAA AAATAAAATG AGTCATTATT TTAAAAAAA AAAAAAAACT  
  
 1946  
 CTAGAGTCTGA CGCAATTG

The sequences of BMP-2 Class I and II, as well as BMP-3 as shown in Tables III, IV, VII and VIII have significant homology to the beta (B) and beta (A) subunits of the inhibins. The inhibins are a family of hormones which are presently being investigated for use in contraception. See, A. J. Mason et al, Nature, 318:659-663 (1985). To a lesser extent they are also homologous to Mullerian inhibiting substance (MIS), a testicular glycoprotein that causes regression of the Mullerian duct during development of the male embryo and transforming growth factor-beta (TGF- $\beta$ ) which can inhibit or stimulate growth of cells or cause them to differentiate. Furthermore, the sequence of Table VII encoding hBMP-2 Class II has significant homology to the Drosophila decapentaplegic (DPP-C) locus transcript. See, J. Massague, Cell, 49:437-438 (1987); R. W. Padgett et al, Nature, 325:81-84 (1987); R. L. Cate et al, Cell 45: 685-698 (1986). It is considered possible therefore that BMP-2 Class II is the human homolog of the protein made from this transcript from this developmental mutant locus.

#### C.. BMP-3

Because bovine and human bone growth factor genes are presumed to be significantly homologous, oligonucleotide probes which have been shown to hybridize to the bovine DNA sequence of Table IV.A and IV.B are used to screen a human genomic library. A human genomic library (Toole et al., supra) is screened using these probes, and presumptive positives are isolated and DNA sequence obtained as described above. Evidence that this recombinant encodes a portion of the human bone inductive factor molecule relies on the bovine/human protein and gene structure homologies.

Once a recombinant bacteriophage containing DNA encoding a portion of the human BMP-3 molecule is obtained the human coding sequence is used as a probe as described in Example V (A) to identify a human cell line or tissue which synthesizes BMP-3. mRNA is selected by oligo (dT) cellulose

chromatography and cDNA is synthesized and cloned in lambda gt10 by established techniques (Toole et al., supra).

Alternatively, the entire gene encoding this human bone inductive factor can be identified and obtained in additional recombinant clones if necessary. Additional recombinants containing further 3' or 5' regions of this human bone inductive factor gene can be obtained by identifying unique DNA sequences at the end(s) of the original clone and using these as probes to rescreen the human genomic library. The gene can then be reassembled in a single plasmid by standard molecular biology techniques and amplified in bacteria. The entire human BMP-3 factor gene can then be transferred to an appropriate expression vector. The expression vector containing the gene is then transfected into a mammalian cell, e.g. monkey COS cells, where the human gene is transcribed and the RNA correctly spliced. Media from the transfected cells are assayed for bone inductive factor activity as described herein as an indication that the gene is complete. mRNA is obtained from these cells and cDNA synthesized from this mRNA source and cloned. The procedures described above may similarly be employed to isolate other species' bone inductive factor of interest by utilizing the bovine bone inductive factor and/or human bone inductive factor as a probe source. Such other species' bone inductive factor may find similar utility in, inter alia, fracture repair.

#### EXAMPLE VI

##### Expression of Bone Inductive Factors.

In order to produce bovine, human or other mammalian bone inductive factors, the DNA encoding it is transferred into an appropriate expression vector and introduced into mammalian cells by conventional genetic engineering techniques.

One skilled in the art can construct mammalian expression vectors by employing the sequence of Tables II-

VIII or other modified sequences and known vectors, such as pCD [Okayama et al., Mol. Cell Biol., 2:161-170 (1982)] and pJL3; pJL4 [Gough et al., EMBO J., 4:645-653 (1985)]. The transformation of these vectors into appropriate host cells can result in expression of osteoinductive factors. One skilled in the art could manipulate the sequences of Tables II-VIII by eliminating or replacing the mammalian regulatory sequences flanking the coding sequence with bacterial sequences to create bacterial vectors for intracellular or extracellular expression by bacterial cells. For example, the coding sequences could be further manipulated (e.g. ligated to other known linkers or modified by deleting non-coding sequences therefrom or altering nucleotides therein by other known techniques). The modified bone inductive factor coding sequence could then be inserted into a known bacterial vector using procedures such as described in T. Taniguchi et al., Proc. Natl Acad. Sci. USA, 77:5230-5233 (1980). This exemplary bacterial vector could then be transformed into bacterial host cells and bone inductive factor expressed thereby. For a strategy for producing extracellular expression of bone inductive factor in bacterial cells., see, e.g. European patent application EPA 177,343.

Similar manipulations can be performed for the construction of an insect vector [See, e.g. procedures described in published European patent application 155,476] for expression in insect cells. A yeast vector could also be constructed employing yeast regulatory sequences for intracellular or extracellular expression of the factors of the present invention by yeast cells. [See, e.g., procedures described in published PCT application WO86/00639 and European patent application EPA 123,289].

A method for producing high levels of an osteoinductive factor of the invention from mammalian cells involves the construction of cells containing multiple copies of the heterologous bone inductive factor gene. The heterologous gene

can be linked to an amplifiable marker, e.g. the dihydrofolate reductase (DHFR) gene for which cells containing increased gene copies can be selected for propagation in increasing concentrations of methotrexate (MTX) according to the procedures of Kaufman and Sharp, J. Mol. Biol., 159:601-629 (1982). This approach can be employed with a number of different cell types.

For example, a plasmid containing a DNA sequence for a bone inductive factor of the invention in operative association with other plasmid sequences enabling expression thereof and the DHFR expression plasmid pAdA26SV(A)3 [Kaufman and Sharp, Mol. Cell. Biol., 2:1304 (1982)] can be co-introduced into DHFR-deficient CHO cells, DUKX-BII, by calcium phosphate coprecipitation and transfection. DHFR expressing transformants are selected for growth in alpha media with dialyzed fetal calf serum, and subsequently selected for amplification by growth in increasing concentrations of MTX (sequential steps in 0.02, 0.2, 1.0 and 5uM MTX) as described in Kaufman et al., Mol Cell Biol., 5:1750 (1983). Transformants are cloned, and biologically active bone inductive factor expression is monitored by rat bone formation assay. Bone inductive factor expression should increase with increasing levels of MTX resistance. Similar procedures can be followed to produce other bone inductive factors.

Alternatively, the human gene is expressed directly, as described above. Active bone inductive factor may be produced in bacteria or yeast cells. However the presently preferred expression system for biologically active recombinant human bone inductive factor is stably transformed CHO cells.

As one specific example, to produce the human bone inductive factor (hBMP-1) of Example V, the insert of U2OS-1 is released from the vector arms by digestion with Sal I and subcloned into the mammalian expression vector pMT2CX digested with Xho I. Plasmid DNA from this subclone is transfected into COS cells by the DEAE-dextran procedure [Sompayrac and

Danna PNAS 78:7575-7578 (1981); Luthman and Magnusson, Nucl. Acids Res. 11: 1295-1308 (1983)]. Serum-free 24 hr. conditioned medium is collected from the cells starting 40-70 hr. post-transfection.

The mammalian expression vector pMT2 Cla-Xho (pMT2 CX) is a derivative of p91023 (b) (Wong et al., Science 228:810-815, 1985) differing from the latter in that it contains the ampicillin resistance gene in place of the tetracycline resistance gene and further contains a XhoI site for insertion of cDNA clones. The functional elements of pMT2 Cla-Xho have been described (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82:689-693) and include the adenovirus VA genes, the SV40 origin of replication including the 72 bp enhancer, the adenovirus major late promoter including a 5' splice site and the majority of the adenovirus tripartite leader sequence present on adenovirus late mRNAs, a 3' splice acceptor site, a DHFR insert, the SV40 early polyadenylation site (SV40), and pBR322 sequences needed for propagation in E. coli.

Plasmid pMT2 Cla-Xho is obtained by EcoRI digestion of pMT2-VWF, which has been deposited with the American Type Culture Collection (ATCC), Rockville, MD (USA) under accession number ATCC 67122. EcoRI digestion excises the cDNA insert present in pMT2-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods. pMT2CX is then constructed by digesting pMT2 with EcoRV and XbaI, treating the digested DNA with Klenow fragment of DNA polymerase I, and ligating Cla linkers (NEBiolabs, CATCGATG). This removes bases 2266 to 2421 starting from the Hind III site near the SV40 origin of replication and enhancer sequences of pMT2. Plasmid DNA is then digested with EcoRI, blunted as above, and ligated to an EcoRI adapter,

5' PO<sub>4</sub>-AATTCCCTCGAGAGCT 3'

3' GGAGCTCTCGA 5'

digested with XhoI, and ligated, yielding pMT2 Cla-Xho, which may then be used to transform *E. coli* to ampicillin resistance. Plasmid pMT2 Cla-Xho DNA may be prepared by conventional methods.

Example VII

Biological Activity of Expressed Bone Inductive Factor

A. BMP-1

To measure the biological activity of the expressed bone inductive factor (hBMP-1) obtained in Example VI above. The factor is partially purified on a Heparin Sepharose column. 4 ml of transfection supernatant from one 100 mm dish is concentrated approximately 10 fold by ultrafiltration on a YM 10 membrane and then dialyzed against 20mM Tris, 0.15 M NaCl, pH 7.4 (starting buffer). This material is then applied to a 1.1 ml Heparin Sepharose column in starting buffer. Unbound proteins are removed by an 8 ml wash of starting buffer, and bound proteins, including BMP-1, are desorbed by a 3-4 ml wash of 20 mM Tris, 2.0 M NaCl, pH 7.4.

The proteins bound by the Heparin column are concentrated approximately 10-fold on a Centricon 10 and the salt reduced by diafiltration with 0.1% trifluoroacetic acid. The appropriate amount of this solution is mixed with 20 mg of rat matrix and then assayed for *in vivo* bone and cartilage formation as previously described in Example III. A mock transfection supernatant fractionation is used as a control.

The implants containing rat matrix to which specific amounts of human BMP-1 have been added are removed from rats after seven days and processed for histological evaluation. Representative sections from each implant are stained for the presence of new bone mineral with von Kossa and acid fuschin, and for the presence of cartilage-specific matrix formation using toluidine blue. The types of cells present within the section, as well as the extent to which these cells display phenotype are evaluated.

Addition of human BMP-1 to the matrix material resulted in formation of cartilage-like nodules at 7 days post implantation. The chondroblast-type cells were recognizable by shape and expression of metachromatic matrix. The amount of activity observed for human BMP-1 was dependent upon the amount of human BMP-1 protein added to the matrix. Table IX illustrates the dose-response relationship of human BMP-1 protein to the amount of bone induction observed.

Table IX

<u>IMPLANT NUMBER</u>	<u>AMOUNT USED</u> (equivalent of ml transfection media)	<u>HISTOLOGICAL SCORE</u>
876-134-1	10 BMP-1	C+2
876-134-2	3 BMP-1	C+1
876-134-3	1 BMP-1	C +/-
876-134-4	10 MOCK	C -
876-134-5	3 MOCK	C -
876-134-6	1 MOCK	C -

Cartilage (c) activity was scored on a scale from 0(-) to 5.

Similar levels of activity are seen in the Heparin Sepharose fractionated COS cell extracts. Partial purification is accomplished in a similar manner as described above except that 6 M urea is included in all the buffers. Further, in a rat bone formation assay as described above, BMP-2 has similarly demonstrated chondrogenic activity.

The procedures described above may be employed to isolate other bone inductive factors of interest by utilizing the bovine bone inductive factors and/or human bone inductive factors as a probe source. Such other bone inductive factors may find similar utility in, inter alia, fracture repair.

The foregoing descriptions detail presently preferred embodiments of the present invention. Numerous modifications

and variations in practice thereof are expected to occur to those skilled in the art upon consideration of these descriptions. Those modifications and variations are believed to be encompassed within the claims appended hereto.

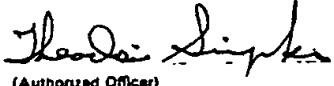
International Application No: PCT/ /

**MICROORGANISMS**Optional Sheet in connection with the microorganism referred to on page \_\_\_\_\_, line \_\_\_\_\_ of the description <sup>1</sup>**A. IDENTIFICATION OF DEPOSIT<sup>1</sup>**Further deposits are identified on an additional sheet  <sup>2</sup>Name of depositary institution <sup>4</sup>

American Type Culture Collection

Address of depositary institution (including postal code and country) <sup>4</sup>12301 Parklawn Drive  
Rockville, Maryland 20852 USA

<u>Name of Deposit</u>	<u>ATCC No.</u>	<u>Referred to on page/line</u>	<u>Date of Deposit</u>
LP-H1	40311	29/20	March 4, 1987
bP50	40295	20/3	December 15, 1986
bP-21	40310	22/18	March 4, 1987
U2OS-3	40342	44/22	June 16, 1987
Lambda U2-OS-1	40343	32/33	June 16, 1987
Lambda BP819	40344	25/23	June 16, 1987
U2OS-39	40345	39/21	June 16, 1987

**C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE<sup>3</sup> (if the indications are not for all designated States)****D. SEPARATE FURNISHING OF INDICATIONS<sup>4</sup> (leave blank if not applicable)**The indications listed below will be submitted to the International Bureau later<sup>5</sup> (Specify the general nature of the indications e.g., "Accession Number of Deposit")E.  This sheet was received with the international application when filed (to be checked by the receiving Office)

  
(Authorized Officer)
 The date of receipt (from the applicant) by the International Bureau <sup>10</sup>

WAS

(Authorized Officer)

## WHAT IS CLAIMED IS:

1.. A pharmaceutical composition comprising a protein selected from the group consisting of:

- (a) BMP-1;
- (b) BMP-2 Class I;
- (c) BMP-2 Class II;
- (d) BMP-3; and

mixtures thereof, in a pharmaceutically acceptable vehicle.

2.. A composition of Claim 1 wherein said protein is BMP-1.

3.. A composition of Claim 1 wherein said protein is BMP-2 Class I.

4.. A composition of Claim 1 wherein said protein is BMP-2 Class II.

5.. A composition of Claim 1 wherein said protein is BMP-3.

6.. The pharmaceutical composition of Claim 1 further comprising a matrix capable of delivering the composition to the site of the bone defect and providing a structure for inducing bone formation.

7.. The composition of Claim 6 wherein said matrix comprises a material selected from the group consisting of hydroxyapatite, collagen, polylactic acid and tricalcium phosphate.

8.. A method for inducing bone formation in a patient in need of same comprising administering to said patient an effective amount of a composition of Claim 1-7.

9.. A process for producing BMP-1 comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-1, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-1 from said culture medium.

10.. A process according to Claim 9 wherein said DNA sequence comprises substantially the nucleotide sequence of Table VI.

11.. A process for producing BMP-2 Class I comprising culturing in a suitable culture medium a cell line transformed

with a DNA sequence encoding BMP-2 Class I, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class I from said culture medium.

12. A process for according to Claim 11 wherein said DNA sequence comprises substantially the nucleotide sequence of Table VII.

13.. A process for producing BMP-2 Class II comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-2 Class II, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class II from said culture medium.

14. A process according to Claim 13 wherein said DNA sequence comprises substantially the nucleotide sequence of Table VIII.

15. A process for producing BMP-3 comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-3, said DNA sequence being in relative association with an expression control sequence therefor and isolating BMP-3 from said culture medium.

16. A process according to Claim 15 wherein said DNA sequence comprises substantially the nucleotide sequence of Table IVA and IVB.

17. A cDNA sequence encoding BMP-1 comprising substantially the nucleotide sequence of Table VI or a sequence which hybridize thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-1.

18. A cDNA sequence encoding BMP-2 Class I comprising substantially the nucleotide sequence of Table VII or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class I.

19. A cDNA sequence encoding BMP-2 Class II comprising

substantially the nucleotide sequence of Table VIII or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class II.

20. A cDNA sequence encoding BMP-3 comprising substantially the nucleotide sequence of Table IVA and IVB or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-3.

**AMENDED CLAIMS**

[received by the International Bureau on 8 December 1987 (08.12.87)  
original claims 6, 8, 10, 12, 14, 16-20 amended;  
new claims 21-23 added; other claims unchanged (13 pages)]

1. A pharmaceutical composition comprising a protein selected from the group consisting of:

- (a) BMP-1;
- (b) BMP-2 Class I;
- (c) BMP-2 Class II;
- (d) BMP-3; and

mixtures thereof, in a pharmaceutically acceptable vehicle.

2.. A composition of Claim 1 wherein said protein is BMP-  
1..

3.. A composition of Claim 1 wherein said protein is BMP-  
2 Class I.

4.. A composition of Claim 1 wherein said protein is BMP-  
2 Class II.

5.. A composition of Claim 1 wherein said protein is BMP-  
3.

6. The pharmaceutical composition of Claim 1 further comprising a matrix capable of delivering the composition to the site of the bone or cartilage defect and providing a structure for inducing bone or cartilage formation.

7. The composition of Claim 6 wherein said matrix comprises a material selected from the group consisting of hydroxyapatite, collagen, polylactic acid and tricalcium phosphate.

8. A method for inducing bone or cartilage formation in a patient in need of same comprising administering to said patient an effective amount of a composition of Claim 1-7.

9. A process for producing BMP-1 comprising the steps of culturing in a suitable culture medium a host cell transformed with a DNA sequence encoding BMP-1, said DNA sequence being in relative association with an expression control sequence therefor; and isolating said BMP-1 therefrom.

10. A process according to Claim 9 wherein said DNA sequence comprises substantially the nucleotide sequence as follows:

10	20	30	50
CTACAGGGCGG CTTCCTCTGC CGCGCGCCCG CCAGC ATG CCC GGC GTG GCC CGC CTG CGG			
			MET Pro Gly Val Ala Arg Leu Pro
65	80	95	110
CTG CTG CTC EGG CTG CTG CTC CGG CGT CCC CGG CGG CTG GAC TTG GGC			
Leu Leu Leu Gly Leu Leu Leu Pro Arg Pro Gly Arg Pro Leu Asp Leu Ala			
125	140	155	
GAC TAC ACC TAT GAC CTG GCG GAG GAG GAC GAC TCG GAG CCC CTC AAC TAC AAA			
Asp Tyr Thr Tyr Asp Leu Ala Glu Glu Asp Asp Ser Glu Pro Leu Asn Tyr Lys			
170	185	200	215
GAC CCC TGC AAG GCG GCT GCC TTT CTT GGG GAC ATT GCC CTG GAC GAA GAG GAC			
Asp Pro Cys Lys Ala Ala Ala Phe Leu Gly Asp Ile Ala Leu Asp Glu Glu Asp			
230	245	260	275
CTG AGG GCC TTC CAG GTA CAG CAG GCT GTG GAT CTC AGA CGG CAC ACA GCT CGT			
Leu Arg Ala Phe Gln Val Gln Ala Val Asp Leu Arg Arg His Thr Ala Arg			
290	305	320	
AAG TCC TCC ATC AAA GCT GCA GTT CCA GGA AAC ACT TCT ACC CCC AGC TGC CAG			
Lys Ser Ser Ile Lys Ala Ala Val Pro Gly Asn Thr Ser Thr Pro Ser Cys Gln			
335	350	365	380
AGC ACC AAC GGG CAG CCT CAG AGG GGA GCC TGT GGG AGA TGG AGA GGT AGA TCC			
Ser Thr Asn Gly Gln Pro Gln Arg Gly Ala Cys Gly Arg Trp Arg Gly Arg Ser			
395	410	425	
CGT AGC CGG CGG CGG ACG TCC CGA CCA GAG CGT GTG TGG CCC GAT GGG GTC			
Arg Ser Arg Arg Ala Ala Thr Ser Arg Pro Glu Arg Val Trp Pro Asp Gly Val			
440	455	470	485
ATC CCC TTT GTC ATT GGG GGA AAC TTC ACT GGT AGC CAG AGG GCA GTC TTC CGG			
Ile Pro Phe Val Ile Gly Gly Asn Phe Thr Gly Ser Gln Arg Ala Val Phe Arg			
500	515	530	545
CAG GCC ATG AGG CAC TGG GAG AAG CAC ACC TGT GTC ACC TTC CTG GAG CGC ACT			
Gln Ala MET Arg His Trp Glu Lys His Thr Cys Val Thr Phe Leu Glu Arg Thr			

560                    575                    590  
 GAC GAG GAC AGC TAT ATT GTG TTC ACC TAT CGA CCT TGC GGG TGC TGC TCC TAC  
 Asp Glu Asp Ser Tyr Ile Val Phe Thr Tyr Arg Pro Cys Gly Cys Cys Ser Tyr  
  
 605                    620                    635                    650  
 GTG GGT CGC CGC GGC GGG GGC CCC CAG GCC ATC TCC ATC GGC AAG AAC TGT GAC  
 Val Gly Arg Arg Gly Gly Pro Gln Ala Ile Ser Ile Gly Lys Asn Cys Asp  
  
 665                    680                    695  
 AAG TTC GGC ATT GTG GTC CAC GAG CTG GGC CAC GTC GTC GGC TTC TGG CAC GAA  
 Lys Phe Gly Ile Val Val His Glu Leu Gly His Val Val Gly Phe Trp His Glu  
  
 710                    725                    740                    755  
 CAC ACT CGG CCA GAC CGG GAC CGC CAC GTT TCC ATC GTT CGT GAG AAC ATC CAG  
 His Thr Arg Pro Asp Arg Asp Arg His Val Ser Ile Val Arg Glu Asn Ile Gln  
  
 770                    785                    800                    815  
 CCA GGG CAG GAG TAT AAC TTC CTG AAG ATG GAG CCT CAG GAG GTG GAG TOC CTG  
 Pro Gly Gln Glu Tyr Asn Phe Leu Lys MET Glu Pro Gln Glu Val Glu Ser Leu  
  
 830                    845                    860  
 GGG GAG ACC TAT GAC TTC GAC AGC ATC ATG CAT TAC GCT CGG AAC ACA TTC TCC  
 Gly Glu Thr Tyr Asp Phe Asp Ser Ile MET His Tyr Ala Arg Asn Thr Phe Ser  
  
 875                    890                    905                    920  
 AGG CGC ATC TTC CTG GAT ACC ATT GTC CCC AAG TAT GAG GTG AAC GGG GTG AAA  
 Arg Gly Ile Phe Leu Asp Thr Ile Val Pro Lys Tyr Glu Val Asn Gly Val Lys  
  
 935                    950                    965  
 CCT CCC ATT GGC CAA AGG ACA CGG CTC AGC AAG GGG GAC ATT GCC CAA GCC CGC  
 Pro Pro Ile Gly Gln Arg Thr Arg Leu Ser Lys Gly Asp Ile Ala Gln Ala Arg  
  
 980                    995                    1010                    1025  
 AAG CTT TAC AAG TGC CCA GCC TGT CGA GAG ACC CTG CAA GAC ACC ACA GGC AAC  
 Lys Leu Tyr Lys Cys Pro Ala Cys Gly Glu Thr Leu Gln Asp Ser Thr Gly Asn  
  
 1040                    1055                    1070                    1085  
 TTC TCC CCT GAA TAC CCC AAT GGC TAC TCT GCT CAC ATG CAC TGC GTG TGG  
 Phe Ser Ser Pro Glu Tyr Pro Asn Gly Tyr Ser Ala His MET His Cys Val Trp  
  
 1100                    1115                    1130  
 CGC ATC TCT GTC ACA CCC GGG GAG AAG ATC ATC CTG AAC TTC AGG TCC CTG GAC  
 Arg Ile Ser Val Thr Pro Gly Glu Lys Ile Ile Leu Asn Phe Thr Ser Leu Asp  
  
 1145                    1160                    1175                    1190  
 CTG TAC CGC AGC CGC CTG TGC TGG TAC GAC TAT GTG GAG GTC CGA GAT GGC TTC  
 Leu Tyr Arg Ser Arg Leu Cys Trp Tyr Asp Tyr Val Glu Val Arg Asp Gly Phe  
  
 1205                    1220                    1235  
 TGG AGG AAG CGG CCC CTC CGA CGC CGC TTC TGC GGG TCC AAA CTC CCT GAG CCT  
 Trp Arg Lys Ala Pro Leu Arg Gly Arg Phe Cys Gly Ser Lys Leu Pro Glu Pro

1250                    1265                    1280                    1295  
 ATC GTC TCC ACT GAC AGC CGC CTC TGG GTT GAA TTC CGC AGC AGC AGC AAT TGG  
 Ile Val Ser Thr Asp Ser Arg Leu Trp Val Glu Phe Arg Ser Ser Ser Asn Trp  
  
 1310                    1325                    1340                    1355  
 GTT GGA AAG GGC TTC TTT GCA GTC TAC GAA GCC ATC TGC GGG GGT GAT GTG AAA  
 Val Gly Lys Gly Phe Ala Val Tyr Glu Ala Ile Cys Gly Gly Asp Val Lys  
  
 1370                    1385                    1400  
 AAG GAC TAT GGC CAC ATT CAA TOG CCC AAC TAC CCA GAC GAT TAC CGG CCC AGC  
 Lys Asp Tyr Gly His Ile Gln Ser Pro Asn Tyr Pro Asp Asp Tyr Arg Pro Ser  
  
 1415                    1430                    1445                    1460  
 AAA GTC TGC ATC TGG CGG ATC CAG GTG TCT GAG GGC TTC CAC GTG GGC CTC ACA  
 Lys Val Cys Ile Trp Arg Ile Gln Val Ser Glu Gly Phe His Val Gly Leu Thr  
  
 1475                    1490                    1505  
 TIC CAG TCC TTT GAG ATT GAG CGC CAC GAC AGC TGT GGC TAC GAC TAT CTG GAG  
 Phe Gln Ser Phe Glu Ile Glu Arg His Asp Ser Cys Ala Tyr Asp Tyr Leu Glu  
  
 1520                    1535                    1550                    1565  
 GTG CCC GAC GGG CAC AGT GAG AGC AGC ACC CTC ATC GGG CGC TAC TGT GGC TAT  
 Val Arg Asp Gly His Ser Glu Ser Ser Thr Leu Ile Gly Arg Tyr Cys Gly Tyr  
  
 1580                    1595                    1610                    1625  
 GAG AAG CCT GAT GAC ATC AAG AGC ACG TCC AGC CGC CTC TGG CTC AAG TTC GTC  
 Glu Lys Pro Asp Asp Ile Lys Ser Thr Ser Arg Leu Trp Leu Lys Phe Val  
  
 1640                    1655                    1670  
 TCT GAC GGG TCC ATT AAC AAA GCG GGC TTT GCC GTC AAC TTT TTC AAA GAG GTG  
 Ser Asp Gly Ser Ile Asn Lys Ala Gly Phe Ala Val Asn Phe Phe Lys Glu Val  
  
 1685                    1700                    1715                    1730  
 GAC GAG TGC TCT CGG CCC AAC CGC CGG GGC TGT GAG CAG CGG TGC CTC AAC ACC  
 Asp Glu Cys Ser Arg Pro Asn Arg Gly Gly Cys Glu Gln Arg Cys Leu Asn Thr  
  
 1745                    1760                    1775  
 CTG GGC AGC TAC AAG TGC AGC TGT GAC CCC GGG TAC GAG CTG GCC CCA GAC AAG  
 Leu Gly Ser Tyr Lys Cys Ser Cys Asp Pro Gly Tyr Glu Leu Ala Pro Asp Lys  
  
 1790                    1805                    1820                    1835  
 CGC CGC TGT GAG GCT GCT TGT GGC GGA TTC CTC ACC AAC AAG CTC AAC GGC TCC ATC  
 Arg Arg Cys Glu Ala Ala Cys Gly Gly Phe Leu Thr Lys Leu Asn Gly Ser Ile  
  
 1850                    1865                    1880                    1895  
 ACC AGC CGG GGC TGG CCC AAG GAG TAC CCC CCC AAC AAG AAC TGC ATC TGG CAG  
 Thr Ser Pro Gly Trp Pro Lys Glu Tyr Pro Pro Asn Lys Asn Cys Ile Trp Gln  
  
 1910                    1925                    1940  
 CTG GTG GCC CCC ACC CAG TAC CGC ATC TCC CTG CAG TTT GAC TTC TTT GAG ACA  
 Leu Val Ala Pro Thr Gin Tyr Arg Ile Ser Leu Gln Phe Asp Phe Glu Thr

1955	1970	1985	2000			
GAG GGC AAT GAT GTG TGC AAG TAC GAC TTC GIG GAG GTG CGC AGT GGA CTC ACA						
Glu Gly Asn Asp Val Cys Lys Tyr Asp Phe Val Glu Val Arg Ser Gly Leu Thr						
2015	2030	2045				
GCT GAC TCC AAG CTG CAT GGC AAG TTC TGT GGT TCT GAG AAG CCC GAG GTC ATC						
Ala Asp Ser Lys Leu His Gly Lys Phe Cys Gly Ser Glu Lys Pro Glu Val Ile						
2060	2075	2090	2105			
ACC TCC CAG TAC AAC AAC ATG CGC GTG GAG TTC AAG TCC GAC AAC ACC GTG TCC						
Thr Ser Gln Tyr Asn Asn MET Arg Val Glu Phe Lys Ser Asp Asn Thr Val Ser						
2120	2135	2150	2165			
AAA AAG GGC TTC AAG GCC CAC TTC TCA GAA AAG AGG CCA GCT CTG CAG CCC						
Lys Lys Gly Phe Lys Ala His Phe Phe Ser Glu Lys Arg Pro Ala Leu Gln Pro						
2180	2195	2210				
CCT OGG GGA CGC CCC CAC CAG CTC AAA TTC CGA GIG CAG AAA AGA AAC OGG ACC						
Pro Arg Gly Arg Pro His Gln Leu Lys Phe Arg Val Gln Lys Arg Pro Asn Arg Thr						
2225	2235	2245	2255	2265	2275	2285
CCC CAG TGAGGCCCTGC CAGGCTCOCC GGACOCCTTG TTACTCAGGA ACCTCACCTT GGAGGAAATG						
Pro Gln						
2295	2305	2315	2325	2335	2345	2355
GGATGGGGGC TTCCGGTGCCCC ACCAACCCCCC CACCTCCACT CTGCCATTCC GGCCCCACCTC OCTCTGGCG						
2365	2375	2385	2395	2405	2415	2425
GACAGAACCTG GTGCTCTCTT CTCCCCACTG TGCCCGTGG OGGACOGGGG ACCCTTCCCCC GTGCGCTTAC						
2435	2445	2455	2465	2475	2485	2495
CCCTCCATT TTGATGGGTTT CTGIGACATT TCCIGITGIG AAGTAAAGA GGGACCCCCIG CGTCCCTGCCT						

**CTAGA**

11. A process for producing BMP-2 Class I comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-2 Class I, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class I from said culture medium.

12. A process for according to Claim 11 wherein said DNA sequence comprises substantially the nucleotide sequence as

follows:

10	20	30	40	50	60	70
GTCGACTCTA GAGIGIGIGT CAGCAGCTGG CTGGGGACTT CTGAACTTG CAGGGAGAAAT AACCTGGCGCA						
80	90	100	110	120	130	140
CCCCACITTG CGGGGGTGOC TTGGCCCGAG CGGAGCCIGC TTGGCCATCT CGGAGCCCA CGGCCCCCTGC						
150	160	170	180	190	200	210
ACTCTCTGGC CTGGCCCGAC ACTGAGAOGC TGTGCCCCAGC GTGAAAAGAG AGACCTGGCG AGGGCACCC						
220	230	240	250	260	270	280
GGGAGAAGGA GGAGGCCAAAG AAAAGGAAAG GACAITCGGT CCTGGCGOCA GGTCCTTTGA CCAGAGTTTT						
290	300	310	320	330	340	350
TCCAATGGGA CGCCTTTCA ATGGACGIGT CGGGGGTGOC TTGGAGAOG GACCTGGGTC TOCTAAAGGT						
370			385			400
CGACCATG GTG GCC GGG ACC CCC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC						
MET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Val						
415	430	445				
CTC CTG GGC GGC GCG GCT GGC CTC GTT CGG GAG CTG CCC CGC AGG AAG TTC CGG						
Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe Ala						
460	475	490	505			
GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC CTG AGC GAG						
Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val Leu Ser Glu						
520	535	550	565			
TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG AGA CCC ACC CCC AGC						
Phe Glu Leu Arg Leu Leu Ser MET Phe Gly Leu Lys Gln Arg Pro Thr Pro Ser						
580	595	610				
AGG GAC GGC GTG GTG CGC CCC TAC ATG CTA GAC CTG TAT CGC AGG CAC TCG GGT						
Arg Asp Ala Val Val Pro Pro Tyr MET Leu Asp Leu Tyr Arg Arg His Ser Gly						
625	640	655	670			
CAG CGG GGC TCA CCC GGC CCA GAC CAC CGG TTG GAG AGG GCA GCC AGC CGA GGC						
Gln Pro Gly Ser Pro Ala Pro Asp His Arg Leu Glu Arg Ala Ala Ser Arg Ala						
685	700	715				
AAC ACT GTG CGC AGC TTC CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG						
Asn Thr Val Arg Ser Phe His His Glu Ser Leu Glu Glu Leu Pro Glu Thr						
730	745	760	775			
AGT GGG AAA ACA ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG						
Ser Gly Lys Thr Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu						
790	805	820	835			
GAG TTT ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT						
Glu Phe Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln MET Gln Asp Ala						

850                    865                    880

TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA AAA  
Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile Lys

895                    910                    925                    940

CCT GCA ACA GCC AAC TOG AAA TTC CCC GTG ACC AGT CTT TTG GAC ACC AGG TTG  
Pro Ala Thr Ala Asn Ser Iys Phe Pro Val Thr Ser Leu Leu Asp Thr Arg Leu

955                    970                    985

GIG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT GTC ACC CCC GCT GTG ATG  
Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp Val Thr Pro Ala Val MET

1000                  1015                  1030                  1045

CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT GGA TTC GTG GTG GAA GTG GCC CAC  
Arg Trp Thr Ala Gln Gly His Ala Asn His Gly Phe Val Val Glu Val Ala His

1060                  1075                  1090                  1105

TTG GAG GAG AAA CAA GGT GTC TCC AAG AGA CAT GGT AGG ATA AGC AGG TCT TTG  
Leu Glu Glu Lys Gln Gly Val Ser Lys Arg His Val Arg Ile Ser Arg Ser Leu

1120                  1135                  1150

CAC CAA GAT GAA CAC AGC TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC  
His Gln Asp Glu His Ser Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly

1165                  1180                  1195                  1210

CAT GAT GGA AAA GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC  
His Asp Gly Lys Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His

1225                  1240                  1255

AAA CAG CGG AAA CGC CIT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC  
Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp

1270                  1285                  1300                  1315

TTC AGT GAC GIG GGG TGG AAT GAC TGG ATT GTG GCT CCG CGG TAT CAC GCC  
Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala

1330                  1345                  1360                  1375

TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG AAC TCC ACT  
Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr

1390                  1405                  1420

AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GGT AAC TCT AAG ATT CCT AAG  
Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys

1435                  1450                  1465                  1480

GCA TGC TGT GTC CGG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CCT GAC GAG  
Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu

1495                    1510                    1525  
AAT GAA AAG GGT GTA TTA AAG AAC TAT CAG GAC ATG GGT GIG GAG GGT TGT GGG  
Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly  
1540                    1553                    1563                    1573                    1583                    1593                    1603  
TGT CGC TAGTACAGCA AAATTAATAA CATAAAATATA TATATATATA TATATTTTAG AAAAAAGAAA  
Cys Arg  
AAAA

13. A process for producing BMP-2 Class II comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-2 Class II, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class II from said culture medium.

14. A process according to Claim 13 wherein said DNA sequence comprises substantially the nucleotide sequence as follows:

10	20	30	40	50	60	70
CCTCTAGAGGG	CAGAGGAGGA	GGGAGGGAGG	GAAGGGAGGC	GGAGGCGCGC	COGGAAGCTA	GGTGAGTGTC
80	90	100	110	120	130	140
GCATCOGAGC	TGAGGGAGGC	GACCCCTGAGA	OGCCCGCTGCT	GCTCOGGCTG	AGTATCTAGC	TTCGTCCTCCC
150	160	170	180	190	200	210
GATGGGATTG	CGCTCCAAGC	TATCTCGAGC	CTGGAGGCGC	ACAGTCCCCG	GCCCTCGGCC	AGGTTCACIG
220	230	240	250	260	270	280
CAACCGTCA	GAGGTCCCCA	GGAGCTGGTG	CTGGGAGGCC	CGCTACTGCA	GGGACCTAIG	GAGCCATTCC
290	300	310	320	330	340	350
GTAGTGCCAT	CCCGAGCAAC	GCACCTGCTGC	AGCTTCCCCG	AGCCTTTCCA	GCAAGTTTGT	TCAAGATTCG
360	370	380	390	400		
CCTGCAAGAA	TCATGGACTG	TTATTTATAIG	CCCTTGTTTC	TGTCAAGACA	CC ATG ATT CCT	
					MET Ile Pro	
417	432		447		462	
GGT AAC CGA ATG CTG ATG GTC GGT TTA TTA TGC CAA GTC CTG CTA GGA GGC GCG						
Gly Asn Arg MET Leu MET Val Val Leu Leu Cys Gln Val Ile Leu Gly Gly Ala						
477		492		507		
AGC CAT GCT AGT TTG ATA CCT GAG ACG GGG AAG AAA AAA GTC GCC GAG ATT CAG						
Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys Lys Lys Val Ala Glu Ile Gln						

522                    537                    552                    567  
 GGC CAC GCG GGA GGA CGC OGC TCA GGG CAG AGC CAT GAG CTC CTG CGG GAC TTC  
 Gly His Ala Gly Gly Arg Arg Ser Gly Gln Ser His Glu Leu Leu Arg Asp Phe  
  
 582                    597                    612                    627  
 GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGC CGC COG CAG CCT AGC AAG  
 Glu Ala Thr Leu Leu Gin MET Phe Gly Leu Arg Arg Pro Gln Pro Ser Lys  
  
 642                    657                    672  
 AGT GCC GTC ATT CGG GAC TAC ATG CGG GAT CTT TAC CGG CTT CAG TCT GGG GAG  
 Ser Ala Val Ile Pro Asp Tyr MET Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu  
  
 687                    702                    717                    732  
 GAG GAG GAA GAG CAG ATC CAC AGC ACT GGT CTT GAG TAT CCT GAG CGC CGG GCC  
 Glu Glu Glu Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala  
  
 747                    762                    777  
 AGC CGG GCC AAC ACC GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC  
 Ser Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile  
  
 792                    807                    822                    837  
 CCA CGG ACC AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC  
 Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile  
  
 852                    867                    882                    897  
 CCT GAG AAC GAG CGG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG GTG  
 Pro Glu Asn Glu Ala Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln Val  
  
 912                    927                    942  
 GAC CAG CGC CCT GAT TGG GAA AGG CGC TTC CAC CGT ATA AAC ATT TAT GAG GTT  
 Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile Tyr Glu Val  
  
 957                    972                    987                    1002  
 ATG AAG CCG CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC ACA CGA CTA CTG GAC  
 MET Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile Thr Arg Leu Leu Asp  
  
 1017                  1032                  1047  
 ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG GAA ACT TTT GAT GTG AGC CCT  
 Thr Arg Leu Val His His Asn Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro  
  
 1062                  1077                  1092                  1107  
 GCG GTC CTT CGC TGG ACC CGG GAG AAG CAG CCA AAC TAT GGG CTA GCC ATT GAG  
 Ala Val Leu Arg Trp Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu  
  
 1122                  1137                  1152                  1167  
 GTG ACT CAC CTC CAT CAG ACT CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC  
 Val Thr His Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser  
  
 1182                  1197                  1212  
 CGA TCG TTA CCT CAA GGG AGT GGG AAT TGG CGC CAG CTC CGG CCC CTC CTG GTC  
 Arg Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val

1227                  1242                  1257                  1272  
 ACC TTT GGC CAT GAT GGC CGG GGC CAT GCC TTG ACC CGA CGC CGG AGG GGC AAG  
 Thr Phe Gly His Asp Gly Arg His Ala Leu Thr Arg Arg Arg Arg Ala Lys

 1287                  1302                  1317  
 CGT AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAT AAG AAC TGC CGG  
 Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys Arg

 1332                  1347                  1362                  1377  
 CGC CAC TCG CTC TAT GIG GAC TTC AGC GAT GTG GGC TGG AAT GAC TGG ATT GTG  
 Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val

 1392                  1407                  1422                  1437  
 GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC TGC CCC TTT CCA CTG  
 Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp Cys Pro Phe Pro Leu

 1452                  1467                  1482  
 GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT GIG CAG ACC CTG GTC AAT TCT  
 Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser

 1497                  1512                  1527                  1542  
 GTC AAT TCC AGT ATC CCC AAA GCC TGT TGT GTG CCC ACT GAA CTG AGT GGC ATC  
 Val Asn Ser Ser Ile Pro Lys Ala Cys Val Pro Thr Glu Leu Ser Ala Ile

 1557                  1572                  1587  
 TCC ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG  
 Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu

 1602                  1617                  1636                  1646                  1656                  1666  
 ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG ATATAACAC  
 MET Val Val Glu Gly Cys Gly Cys Arg

 1676                  1686                  1696                  1706                  1716                  1726                  1736  
 CACACACACA CACCACATAC ACCACACACA CAACGTTCCA TOCACCTACC CACACACTAC ACAGACTGCT

 1746                  1756                  1766                  1776                  1786                  1796                  1806  
 TCCCTATAGC TGGACCTTAA TTTAAAAAAA AAAAAAAA AATGGAAAAA ATCCCTAAAC ATTCACTTGT

 1816                  1826                  1836                  1846                  1856                  1866                  1876  
 ACCCTATTAA TGACCTTAAG TGCAAATGTT TTGACCATAT TGATCATATA TTTTGACAAA ATATAATTAT

 1886                  1896                  1906                  1916                  1926                  1936                  1946  
 AACCTAOGTAT TAAAAGAAAA AAATAAAAATG AGTCATTATT TTAAAAAAA AAAAAAAACT CTAGAGTOGA

 CGGAATTG

15. A process for producing BMP-3 comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-3, said DNA sequence being in relative

association with an expression control sequence therefor and isolating BMP-3 from said culture medium.

16. A process according to Claim 15 wherein said DNA sequence comprises substantially the nucleotide sequence as follows:

383	393	403	413	428
GAGGAGGAAG CGGTCTAOGG GGGTCCCTCT GCCTCTGCAG AAC AAT GAG CCT CCT GGG GCA				
				Asn Asn Glu Leu Pro Gly Ala
443		458		473
GAA TAT CAG TAC AAG GAG GAT GAA GTA TGG GAG GAG AAG CCT TAC AAG ACT				
		Glu Tyr Gln Tyr Lys Glu Asp Glu Val Trp Glu Arg Lys Pro Tyr Lys Thr		
503		518		533
CIT CAG ACT CAG CCC CCT GAT AAG AGT AAG AAC AAA AAG AAA CAG AGG AAG GGA				
		Leu Gln Thr Gln Pro Pro Asp Lys Ser Lys Asn Lys Lys Gln Arg Lys Gly		
548		563		578
CCT CAG CAG AAG AGT CAG ACG CTC CAG TTT GAT GAA CAG ACC CTG AAG AAG GCA				
		Pro Gln Gln Lys Ser Gln Thr Leu Gln Phe Asp Glu Gln Thr Leu Lys Lys Ala		
608		623		638
AGA AGA AAG CAA TGG ATT GAA CCC CGG AAT TGT GCC AGA CGG TAC CIT AAA GTG				
		Arg Arg Lys Gln Trp Ile Glu Pro Arg Asn Cys Ala Arg Arg Tyr Leu Lys Val		
653		668		683
GAC TTC CCA GAT ATT GCC TGG AGC GAA TGG ATT ATT TCC CCC AAG TCC TTC GAT				
		Asp Phe Ala Asp Ile Gly Trp Ser Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp		
713		728		743
GCC TAT TAC TGC TCC GGA GCG TGC CAG TTC CCC ATG CCA AAG GTAGCCATTG TTTTTGTCC				
		Ala Tyr Tyr Cys Ser Gly Ala Cys Gln Phe Pro MET Pro Lys		
776		786		
TGTCCCTCOCC ATTTCOCATAG ; and				
284	294	304	319	
CTAACCTGIG TTCTCCCTTT TOGTTCTTAG TCT TTG AAG CCA TCA AAT CAC GCT ACC				
		Ser Leu Lys Pro Ser Asn His Ala Thr		
334	349	364	379	
ATC CAG AGT ATA GTG AGA GCT GTG GGG GTC GTC CCT GGA ATC CCC GAG CCT TGC				
		Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile Pro Glu Pro Cys		
394		409		424
TGT GTG CCA GAA AAG ATG TCC TCA CTC AGC ATC TTA TTC TTT GAT GAA AAC AAG				
		Cys Val Pro Glu Lys MET Ser Ser Leu Ser Ile Leu Phe Phe Asp Glu Asn Lys		
439				

454                    469                    484  
AAT GTG GTA CTT AAA GTC TAT CCA AAC ATG ACA GTC GAG TCT TGT GCT TGC AGA  
Asn Val Val Leu Lys Val Tyr Pro Asn MET Thr Val Glu Ser Cys Ala Cys Arg  
  
503                513                523                533  
TAACCTGGTG AAGAACTCAT CTGGATGCTT AACCTCAATCG.

17. A cDNA sequence encoding BMP-1 comprising substantially the nucleotide sequence recited in Claim 10 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-1.

18. A cDNA sequence encoding BMP-2 Class I comprising substantially the nucleotide sequence recited in Claim 12 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class I.

19. A cDNA sequence encoding BMP-2 Class II comprising substantially the nucleotide sequence recited in Claim 14 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class II.

20. A cDNA sequence encoding BMP-3 comprising substantially the nucleotide sequence recited in Claim 16 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-3.

21. A vector containing a DNA sequence encoding an osteoinductive protein and heterologous DNA, the DNA sequence encoding the protein being selected from the group consisting of:  
a. a DNA sequence encoding BMP-1 comprising substantially the nucleotide sequence recited in Claim 10 or a sequence which

hybridize thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-1;

b. a DNA sequence encoding BMP-2 Class I comprising substantially the nucleotide sequence recited in Claim 12 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class I;

c. a DNA sequence encoding BMP-2 Class II comprising substantially the nucleotide sequence recited in Claim 14 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class II; and

d. a DNA sequence encoding BMP-3 comprising substantially the nucleotide sequence recited in Claim 16 or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-3.

22. A cell transformed with a vector according to claim 21 which is capable of expressing a DNA sequence encoding the osteoinductive protein and progeny of said cell.

23. The transformed cell according to claim 24 selected from the group consisting of a mammalian cell, a bacterial cell, an insect cell, and a yeast cell.

## **INTERNATIONAL SEARCH REPORT**

International Application No PCT/US87/01537

**I. CLASSIFICATION OF SUBJECT MATTER** (if several classification symbols apply, indicate all) <sup>3</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC  
IPC(4): C07K 13/00,15/00; A61K 37/00; See Attachment  
US CL: 530/350,395,397; 514/12; 536/27 See Attachment

## **II. FIELDS SEARCHED**

Classification System   Classification Symbols	
US	530/350, 395, 397; 514/12; 536/27 435/68, 70, 172.3; 935/13

**Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched**

COMPUTER SEARCH CAS, APS: BONE MORPHOGEN, BONE  
INDUCTIVE PROTEIN, BMP, OSTEOINDUCTIVE FACTOR

### III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>14</sup>

Category	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
<u>X, P</u>	US, A, 4,619,989 (URIST) 28 Oct 1986.	1-8 9-20
<u>Y, P</u>		
<u>X</u>	US, A, 4,563,350 (NATHAN ET AL)	1-8
<u>Y</u>	7 January 1986.	9-20
<u>X</u>	US, A, 4,455,256 (URIST) 19 June 1984.	1-8
<u>Y</u>		9-20
<u>X</u>		
<u>Y</u>		
<u>—</u>		
<u>—</u>		
<u>X</u>	Proc. Natl. Acad. Sci USA, Vol. 81,	1
<u>—</u>		

- **Special categories of cited documents:** i)  
“A” document defining the general state of the art which is not considered to be of particular relevance  
“E” earlier document but published on or after the international filing date  
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
“O” document referring to an oral disclosure, use, exhibition or other means  
“P” document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
  - "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
  - "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
  - "&" document member of the same patent family

---

**IV. CERTIFICATION**

**Date of the Actual Completion of the International Search :**

Date of Mailing of this International Search Report:

08 October 1987

20 OCT 1987

**International Searching Authority:**

ISA/US

Signature of Authorized Officer <sup>to</sup>  
*Alvin E. Tanenholtz*  
ALVIN E. TANENHOLTZ

PCT/US87/01537

Attachment To Form PCT/ISA/210, Part I.

IPC(4): C12P 21/00, 21/02; C12N 15/00; C07H 15/12

US CL : 435/68, 70, 172.3; 935/13

## III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No <sup>18</sup>
Y	<u>Science</u> , Vol. 220 issued 13 May 1983 (Washington, D.C.) (URIST) "Bone cell Differentiation and Growth Factors" pages 680-686.	1-20
Y	<u>Proc. Natl. Acad. Sci., USA</u> , Vol. 80 issued November 1983 (Washington, D.C.) (SAMPATH ET AL), "Homology of bone- inductive proteins from human monkey, bovine and rat extracellular matrix," pages 6591-6595.	1-20
Y	<u>Proc. Natl. Acad. Sci., USA</u> , Vol. 78 issued November 1981, (Washington, D.C.) (SUGGS ET AL), "Use of synthetic oligonucleotides as hybridization probes: Isolation of cloned cDNA sequence for human $\beta_2$ -microglobulin" pages 6613-6617.	1-20